Cardiff University

School of Bioscience

Protein Engineering Utilising Single Amino Acid Deletions Within *Photinus Pyralis* Firefly Luciferase

Lisa M. Halliwell

September 2015

Cardiff School of Biosciences

The Sir Martin Evans Building

Museum Avenue

Cardiff

CF10 3AX

A thesis submitted for the degree of Doctor of Philosophy for Cardiff University

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed ..... (candidate) Date 15/09/2015

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of ...PhD.....(insert MCh, MD, MPhil, PhD etc, as appropriate)

Signed ..... (candidate) Date 15/09/2015.....

STATEMENT 2

This thesis is the result of my own independent work/investigation, except where otherwise stated.
Other sources are acknowledged by explicit references.

Signed ..... (candidate) Date 15/09/2015.....

STATEMENT 3

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ..... (candidate) Date 15/09/2015....

STATEMENT 4: PREVIOUSLY APPROVED BAR ON ACCESS

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loans **after expiry of a bar on access previously approved by the Graduate Development Committee.**

Signed ..... (candidate) Date 15/09/2015....

STATEMENT 5

Work involving human participants and/or samples was performed with the approval of the appropriate Research Ethics Committee.

Signed ..... (candidate) Date 15/09/2015...

Abstract

The bioluminescence reaction is catalysed by firefly luciferase, converting the substrates D-luciferin, ATP and molecular oxygen with Mg^{2+} to produce light and this reaction has had wide ranging implications within a number of fields from industry to academia. The discovery of luciferase has been revolutionary in the real time *in vivo* study of cells given that it requires no energy for excitation, delivering a high signal to background ratio providing a highly sensitive assay. This protein, to date, has been utilised in molecular cell biology, cellular imaging, microbiology and numerous other fields. The extensive application of this protein has paved the way for the generation of toolbox of variants with altered properties. Protein engineering involving substitution mutations made within *Photinus pyralis* (*Ppy*) FLuc has led to the discovery of a number of novel variants however there is a bank of growing evidence displaying the power of deletions as an alternative for the development of proteins with altered properties since deletions can sample structural diversity not accessible to substitutions alone.

A novel mutagenic strategy was implemented to incorporate single amino acid deletions within thermostable firefly luciferase (x11FLuc) targeting loop structures (M1-G10, L172-T191, T352-F368, D375-R387, D520-L526, K543-L550). Of 43 deletion mutants obtained, 41 retained bioluminescent activity and other characteristics such as resistance to thermal inactivation. Surprisingly, only 2 variants, ΔV365 and ΔV366, exhibited a complete loss of activity showing that the luciferase protein is largely tolerant to single amino acid deletions.

In order to identify useful mutants in the extensive library, a 96-well format luminometric cell lysate assay was developed which indicated the effect of deletions was largely region specific, for example, N- terminal deletions did not alter the activity of x11FLuc, whilst deletions within L172- T191, D375-R387, D520-L526 and the C- terminal loop reduced overall activity. On the other hand, deletions within T352-F368 enhanced overall bioluminescent activity and remarkably exhibited other important characteristics such as increases in specific activity and a reduced K_M for luciferin. Therefore, a novel motif (omega loop) was identified as important for FLuc activity after full characterization of mutants.

Characterisation of the deletion mutants originating from the omega loop (T352-F368), ΔP359 and ΔG360 both presented a reduced K_M for luciferin, whilst ΔA361, ΔV362, ΔG363 presented an increase in K_M towards ATP as compared to x11FLuc. Thus, deletions in the omega loop, in the main, improved activity and altered reaction kinetics, in particular ΔG363 retained 53% of initial activity after 250s. As such, it is considered that the field of protein engineering should not only overlook the utility of single amino acid deletions, since such mutagenic strategies may sample structural space not achieved by substitutions alone and mutations within less popularized secondary structures such as omega loops are can act as a useful tool in the improvement of proteins.

Acknowledgements

I would like to start by saying a thank you to the research group in which I have been a part, in particular, Angela Marchbank, and Joanne Kilby, whom throughout this past year have gone out of their way to support this project be it through answering my endless technical questions to ensuring that the materials needed throughout were always in stock and available.

A special thank you must go to my mentor, Amit Jathoul, without whose help this thesis and the data within would not have been possible. I am so grateful for his consistent patience, encouragement and commitment to this project and I have found his enthusiasm and overwhelming knowledge on the field of bioluminescence and luciferase protein engineering to be both inspiring and a constant source of motivation.

Another special thank you must go to my supervisor, Prof Jim Murray. Thank you for allowing me to explore a topic in which I have always had a keen interest but more than that, thank you for the continued guidance, support and understanding you have provided me over the last 4 years.

Lastly, thank you to my wonderful family and my equally wonderful boyfriend. Thank you for the putting up with the mood swings, for providing never ending cups of coffee and for just being there when I felt that this thesis would not be possible. Thank you for your endless support over the completion of this work and as such, this is for you.

Table of Contents

Common Abbreviations

ATP	Adenosine triphosphate
CIEEL	Chemically-Induced Electron Exchange Luminescence
CoA	Coenzyme A
d.p.	Decimal place
FF	Firefly
FWHM	Full width half maximum
Imax	Maximum observed intensity
IMD	Imidazole
IPTG	Isopropylthiogalactoside
KO	Knock out
L.AMP	Dehydroluciferyl-adenylate
LB	Luria Bertani medium
LH_2	Beetle D-luciferin
LH_2-ATP	Luciferyl-adenylate
LO	Oxyluciferin
LO.AMP	Oxyluciferyl-adenylate
Luc	Beetle luciferase
MM	Michaelis-Menten
PCR	Polymerase chain reaction
PMT	Photomultipler tube
PPi	Pyrophosphate
Ppy	*Photinus pyralis*
SDM	Site directed mutagenesis
UG	Ultra-glo ^{TM} luciferase
WT	Wild-type

"Some things, indeed, are not seen in daylight, though they produce sensation in the dark: as for example the things of fiery and glittering appearance for which there is no distinguishing name, " Aristotle De Anima (BookII, Chap 7, Sec4), In R.D. Hicks translation

Chapter 1
Introduction

1.1. Bioluminescence

Humans have always irrevocably been fascinated by organisms with "electric organs" and indeed from Pliny to Aristotle to Darwin all have made commentary about these curiosities of nature, those organisms with the characteristic of 'self-luminosity'.

Bioluminescence, as coined by Harvey (1940) and literally translated as "living light", has been noted to be widely distributed throughout many different ecosystems, from those that walk on the land to those that inhabit the lower reaches of the ocean where bioluminescence is no longer considered as the exception but the rule (Harvey, 1940). Simply, bioluminescence is a natural form of chemiluminescence through which living organisms convert chemical energy into the emission of light and it was in 1885, that the German scientist, Emil du Bois-Reymond, (McElroy and Hastings, 1955) identified two extracts from jellyfish which were responsible for this striking phenomenon when mixed together, one heat sensitive and one heat resistant which over the next 70 decades were termed luciferase (Luc) and luciferin respectively (Green and McElroy, 1956).

1.1.1. Bioluminescent Systems

1.1.1.1. The Evolution of Bioluminescent Systems

Bioluminescence is noted throughout the natural world, from bacteria to insects and marine life, and, at present, there are known to be some 30 differing bioluminescent systems which whilst all displaying bioluminescence, utilise differing bioluminogenic substrates spanning a number chemical classes (Wilson and Hastings, 1998; Purtov *et al.*,

2015). Indeed, between the Lucs, there is large variation with little sequence and structural homology and the Luc proteins vary from between 20kDa to 62kda in size (Wilson and Hastings, 1998). This is indicative that bioluminescence itself arose as a result of convergent evolution, whereby these novel characteristics, arose on multiple occasions from different origins. This necessity for nature to have selected the development of bioluminescent systems is considered amongst five key theories, those being that luminous organs are essential for camouflage, attraction, repulsion, communication and illumination (Haddock *et al*., 2010).

Despite the lack of homology between enzymes, and a range of substrates which are chemically unrelated, all bioluminescent systems are considered as Luc-catalysed reactions involving oxygenation of a bioluminogenic substrate (Hastings and Tu, 1981), and there is some commonality, whereby the decay of a Luc-bound-peroxy-luciferin intermediate provides energy for excitation (Wilson and Hastings, 1998). Due to this, it is considered that bioluminescence initially arose as a function to detoxify reactive oxygen species (Day *et al*., 2004) which is concurrent with the requirement of oxygen to provide adequate energy for emission.

Concurrent with the theory of detoxification, with regards to luciferin, studies report of the antioxidant properties against oxidative and nitrosative stress (Dubuisson *et al*., 2004). It is considered that due to this vital role that bioluminescence arose, not due to Luc, but rather the luciferin as a chemical to remove oxidative species that would otherwise be toxic to cells (Dubuisson *et al*., 2004). Luciferin can be considered to be key to the evolution of bioluminescence, since Luc acts as a catalyst and primarily as a solvent cage to protect the emitter from environmental photophysical or photochemical deactivation processes in the active site of Luc, resulting in higher intensity than would be achieved by chemilluminescence (Vassel *et al*. 2012). It is considered that following the evolution of firefly luciferin (LH_2), the Luc evolved from other metabolic pathways involved in the synthesis of fatty acids.

1.1.1.2. Common Bioluminescent Systems

It is estimated that 80% of bioluminescent systems originate from organisms that inhabit the oceans, whereas 20% may be attributed to terrestrial organisms. Despite this disparity, bioluminescence originating from terrestrial organisms, such as fireflies, has been more heavily investigated and characterized. This is due to the propensity for Lucs derived from terrestrial organisms to have evolved typically brighter luminescence due to the generally higher background or incident light with which they compete (Seliger and McElroy, 1959). Amongst the most studied bioluminescent systems include those that originate within beetles (coleoptera), and amongst marine organisms, the sea pansy, dinoflagellates, copepods and jellyfish. Lucs derived from the oceans tend to emit within a narrower range of wavelengths (circa. 440-478nm) (Shimomura and Johnson, 1975), since emissions within this range confer the greatest optical transparency within the oceans.

The Lucs and the bioluminescent reactions in which they take part are numerous. Lucs derived from the sea pansy (*Renilla reniformis*) and copepods (*Gaussia princeps*), coined RLuc and GLuc respectively, both utilize a bioluminogenic substrate benzylimidazo-pyrazinone coelenterazine in a calcium dependent manner to produce coelenteramide with the concurrent emission of a photon at circa 480nm. On the other hand, whilst many Lucs are found intracellularly, Lucs have been identified which are secreted originating from marine ostracods. Such Lucs include that derived from *Vargula hilgendorfii* and more recently *Cypridina noctiluca* both of which are characteristically very similar (Thompson *et al*., 1990) (Nakajima *et al*., 2004). Many marine bioluminescent systems also employ photoproteins such as aequorin which are coupled to fluorescent proteins (such as the green fluorescent proteins, GFP) to red-shift bioluminescence via the exploitation of bioluminescence resonance energy transfer (BRET) (Shimomura, 1995). It is theorised that such red shifted bioluminescence occurs amongst those marine organisms able to see red light to act as a tool to attract mates and hunt without being detected (Schrope, 2007).

Bioluminescence may also be observed within bacteria which unlike many other bioluminescent systems which require the addition of substrates for bioluminescence, the lux operon encodes all the components necessary to achieve bioluminescence thereby being a wholly contained system (Tu and Mager, 2008). This bacterial system encodes a heterodimeric bacterial luciferase which reduces a further encoded, flavin mononucleotide, the product of which is a peroxyflavin responsible for emission (circa. 490nm) (Tu and Mager, 2008).

Despite the number of bioluminescent systems available, perhaps the most versatile belongs to the beetle Lucs which catalyze the oxidation of a benzothiazole-based LH_2 in a reaction requiring ATP, Luciferin, Mg^{2+} and molecular oxygen and emits within a wide spectral range of ca. 550-630nm (White et al., 1971; Viviani *et al*., 2002).

1.1.2. The Beetle Luciferases

Of the terrestrial organisms, bioluminescence is most prevalent in the three families of luminous beetles, those being the true fireflies (*Lampyridae*), glow worms (*Phengodidae*) and click beetles (*Elateridae*) all of which utilize the same LH_2 molecule (Wood, 1990).

1.1.2.1. Firefly Bioluminescence

The beautiful glow in the North American summer night sky can be attributed, in part, to the lower abdomen of the family of fireflies, Lampyridae, in which resides a structure known as the lantern. Belonging to the order Coleptera, firefly luciferase (FLuc), key to the bioluminescence reaction, resides within this light emitting organ, confined to the peroxisomes of the cells due to the small peptide signal sequence residing within the C terminus (Conti *et al.*, 1996). Bioluminescence remained a mystery for decades since it was first reported by Emil Du Bois and it was only in the 1950s, close to a century after Du Bois, that William McElroy, Emil White and Howard Seilger began to undertake basic research of the Lucs, starting with *Photinus pyralis (Ppy).*

Firefly bioluminescence was the first system to receive extensive investigation with modern biochemical techniques, the most studied of this family being the North American firefly or *Ppy*. More recently, FLucs from other families have been studied due in part, to the interest in elucidating the mechanism underlying the colour differences, as different beetle and even different species of fireflies display bioluminescence of a wide variety of colours. The extent to which the North American FLuc has been studied can be in part attributed to the quantum yield, which defines the number of luciferin molecules required to produce a single photon, which although revised down to 41.0 ±7.4% in 2008 (Ando *et al.,* 2008) and later confirmed Niwa *et al.* (2010), nevertheless remains as one of the most efficient of all known Lucs.

1.1.2.2. Evolution of Firefly Bioluminescence

Key to the evolution of bioluminescence within fireflies, as seen amongst other species, was selection for bioluminescence in communication, reproduction and in the attraction of prey (Haddock *et al.,* 2010). Amongst the fireflies, bioluminescence is heavily implicated in mating behaviors, allowing for flash synchrony, intensity and the flashing patterns characteristic of fireflies (Lloyd, 1983), although it has been noted to be also involved in predator deterrence, whereby lucibufagin compounds in fireflies are distasteful and toxic to predators, who associate this with light emission (Eisner *et al.*, 1978; Knight *et al.,* 1999).

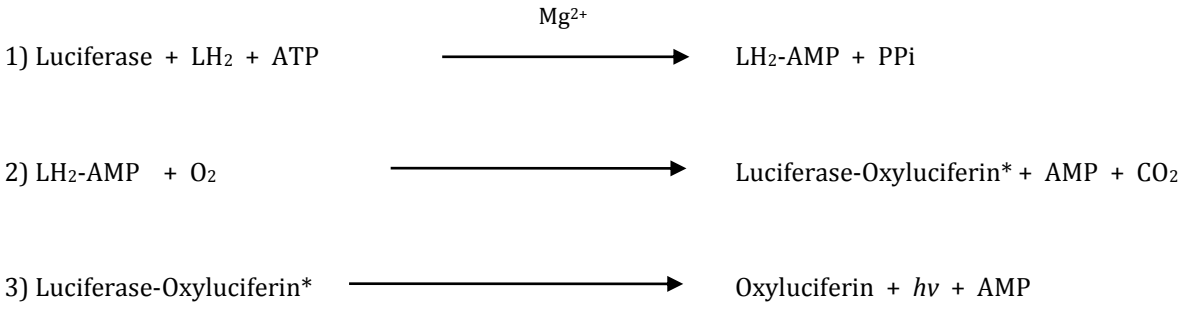
Biochemically, the FLuc has not only been associated with bioluminescence, it has been noted to be implicated in alternative pathways which highlight the potential origins of bioluminescence within these insects. FLuc is homologous to common fatty acyl-CoA synthetase enzymes (Oba *et al*., 2003) and when taken together in combination with phylogentic studies suggests that FLuc arose from this family of super adenylating enzymes via gene duplication (Oba *et al*., 2006). Other enzymes belonging to this family of super adenylating enzymes includes iron-binding siderphores, gramicidin S, tyrocidine, penicillin, fatty acyl CoA ligase and 4-coumarate CoA ligase (Conti *et al*., 1996).

1.1.3. The Coleopteran Bioluminescent Reaction

The core mechanism underlying bioluminescence interested researchers since the initial determination of the components and it was only until McElroy & Deluca in 1978 who proposed that the general scheme of the firefly bioluminescent reaction must be a multistep process, as outlined in equation 1.1, where the production of 'cold light' can be split into two processes utilising the substrates D-LH_2, ATP (adenosine triphosphate), molecular oxygen and a metallic cation, commonly Mg^{2+}.

Beetle luciferase (EC 1.13.12.7), chemically acknowledged as luciferin 4-monooxygenase, is classified as an oxidoreductase enzyme (White *et al*., 1971) and catalyzes the ATP-dependent conversion of the LH_2, 2-(6-hydroxy-2-benzothiazolyl)-2-thiazoline-4-carboxylic acid, into an electronically excited species. Characteristically, oxygenases typically involve the use of a redox prosthetic group and it is interesting to note that this is not such the case with firefly bioluminescent reaction. The core mechanism that underlies this conversion is a Sn2 nucleophillic displacement reaction (Marques and Silva, 2009).

Equation 1.1. Bioluminescent Reaction of Beetle Luciferase

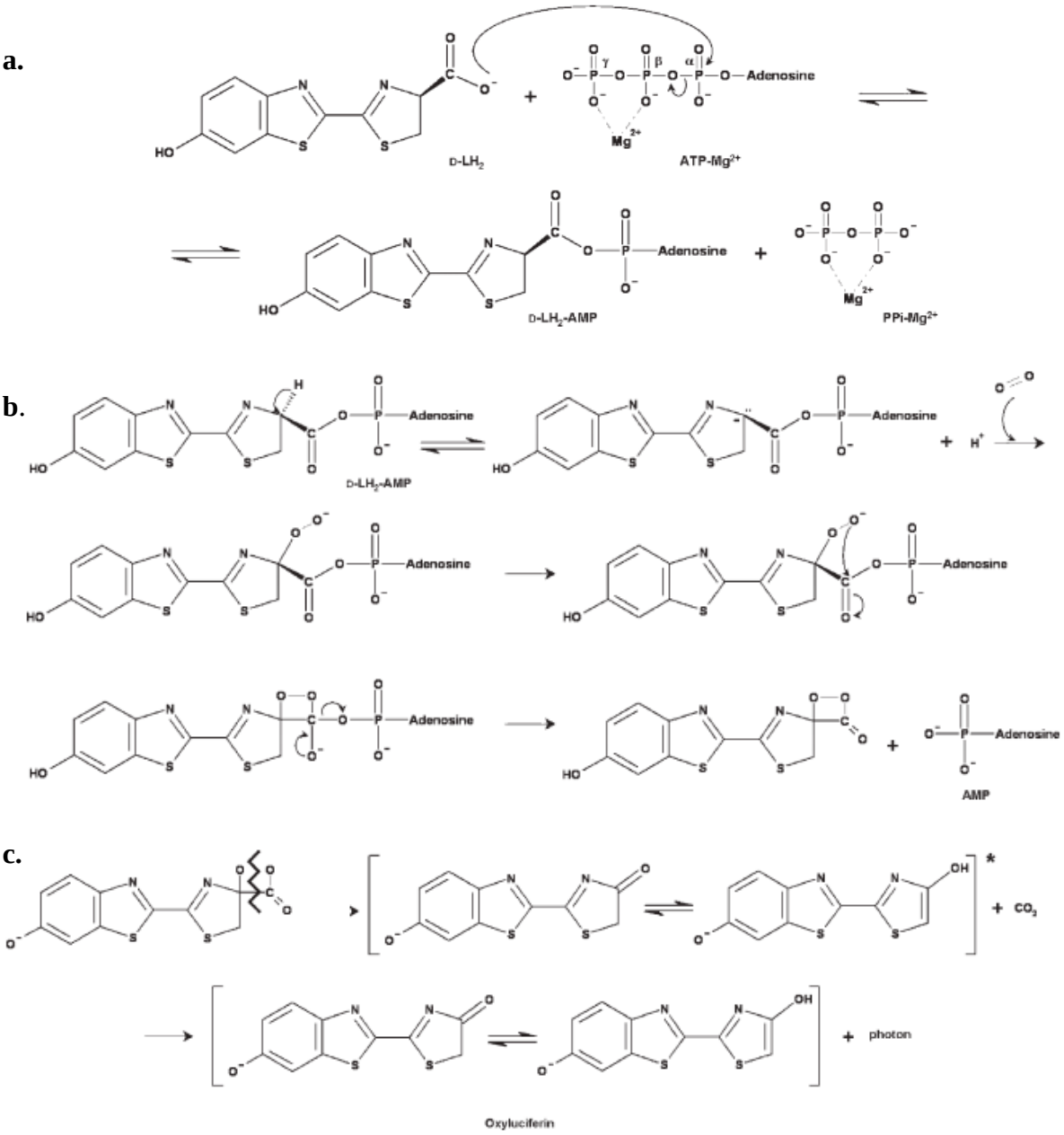


Key: *designates excited species; ATP: adenosine triphosphate; PPi: inorganic pyrophosphate; AMP: adenosine monophosphate; CO_2: carbon dioxide; *hv*: photon

The first step of the enzymatic reaction involves the binding of LH_2 and Mg.ATP with the FLuc, and then an adenylation reaction which gives rise to an enzyme bound luciferyl-adenylate (LH_2-AMP). The oxygen of the carboxyl moiety within the thiazoline ring of LH_2 is a nucleophile and a proton is transferred to the electrophilic moiety α-phosphoryl group of the ATP and it is this transfer that displaces inorganic pyrophosphate and couples the resulting adenylate to the LH_2 (Marques and Silva, 2009).

Following the formation of the LH_2-AMP, the second step involves the incorporation of a single atom of oxygen and the decarboxylation of the LH_2, where the intermediate reacts with oxygen leading to the production of a spin-singlet excited species, oxylucferin, CO_2 and AMP (White *et al*., 1971). It is considered that the loss of a proton at C4 resulting in the formation of a carbanion is a key step in this process and the nature of this loss is conjectured (Marques and Silva, 2009; Branchini *et al*., 2015). Whilst adenylation increases the pKa of the C4 and is considered to be a defining factor for this loss of a proton, it is also considered that an enzymatic base may function to catalyse this removal (Branchini *et al*., 1998). As a carbanion, it is highly reactive and will rapidly undergo nucleophilic attack of molecular oxygen resulting in a hydroperoxide (Marques and Silva, 2009). Subsequently, the hydroperoxide undergoes an internal nucleophilic attack which generates a four membered cyclic intermediate, termed the LH_2 dioextanone, resulting in the displacement of the AMP (Marques and Silva, 2009). Throughout all luminescent reactions, the LH_2 dioextanone is a key intermediate for all bioluminescent emitters since the spontaneous decarboxylation of this energy rich moiety releases the some 0.2kJ/mole (with regards to *Ppy* FLuc) which is concurrent with the energy required to generate the high yield singlet excited state oxyluciferin, CO_2 and AMP (Wilson and Hastings, 1998). The mechanism for this is by a postulated chemically initiated electron exchange luminescence (CIEEL) (Figure 1.1) (McElroy and Hastings, 1955; White *et al*., 1971; Koo *et al*., 1980). Bioluminescence is emitted as the excited oxyluciferin rapidly relaxes to the ground state, with the concurrent release of a photon of light emitting at an λmax circa. 550-560nm corresponding to yellow-green light (McElroy and Seilger, 1961).

Figure 1.1. Chemically Initiated Electron Exchange Luminescence (CIEEL)



Mechanism for CIEEL. a. Adenylation of D-LH_2 via the displacement of PPi from ATP forming D-luciferyl-adenylate**. b**. Proton loss from D-LH_2-AMP leads to the formation of a reactive carbanion able to undergo nucleophillic attack of molecular oxygen. The resulting hydroperoxide undergoes subsequent internal nucleophilic attack to the electrophilic carbon displacing AMP, producing a cyclic dioxethanone ring. **c.** Spontaneous oxidative decarboylation of the dioetanone produces an excited state oxyluciferin able to decay to ground level. (Marques and Silva, 2009).

While some reports suggest that it is only when oxyluciferin is dissociated from the enzyme that bioluminescence may be emitted, it is more commonly believed that the oxyluciferin must remain bound to FLuc during decay (Gandleman *et al*., 2010). This is understood on the premise that a dissociated excited species, free within a solvent, would non-radiatively decay resulting in an absence of light emission.

1.1.4. Structure of Firefly Luciferase

Following the crystallisation of FLuc (Green and McElroy, 1956), the X ray structure was finally solved to a 2.0Å resolution in 1996 (Conti *et al*., 1996). Further study of the enzyme and its structure has greatly expanded our understanding of not only the fold of the protein, but enabled the elucidation of the mechanism of bioluminescence.

FLuc has an unusual fold shared with the acyl-CoA ligases and is grouped into two major domains linked via a four residue loop. The secondary structure of the N- terminal domain lies between 1 and 436 amino acid residues and is composed of a ß barrel, two ß sheets, two ß strands and three α helices (Conti *et al*., 1996). The C- terminal domain on the other hand is the smaller of the two domains and lies between 440-550 amino acid residues and is comprised of five ß-strands and three α-helices (Conti *et al*., 1996) (Figure 1.2.).

The sequence of the gene encoding FLuc, and more specifically FLuc isolated from *Ppy*, contains 1976bp genomic DNA (Wet *et al*., 1987), translating into 550 amino acid residues. The gene itself is shorthanded to *luc* however, in this Thesis, this abbreviation may also refer to the enzyme itself (as Branchini *et al*., 2001).

Generally, proteins have a maturation time during which the polymer of amino acids comprising the proteins folds into the correct functioning structure. In the case of FLuc, this maturation time has been deduced to be small since the first 190 residues fold co-translationally and the rest of the polypeptide folds rapidly on release from the ribosome (Svetlov *et al*., 2007).

1.1.4.1. The Firefly Luciferase Active site and Catalysis

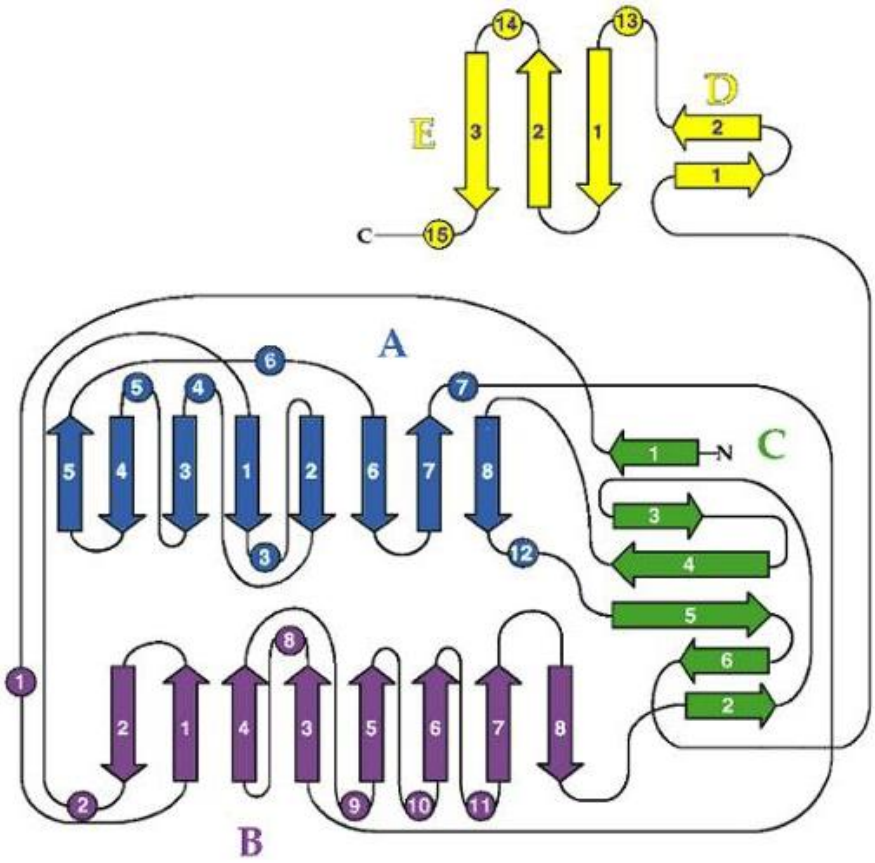
The active site itself, responsible for the binding of the substrate and the co-factors lies between the two domains (Conti et al., 1996). In the unbound state, the two domains are relatively far apart (Conti *et al*., 1996) as such, it is generally considered that due to the small but highly flexible linker between the two domains, once there is binding of the substrate, the protein effectively "sandwiches", bringing the two domains within closer proximity and closing around the substrates (Conti *et al*., 1996) (Figure 1.3). This closed conformational state of the protein holds the benzothiazole ring of the LH_2 in a hydrophobic pocket allowing for deprotonation within the active site of FLuc (Figure 1.3), resulting in higher intensity than chemiluminescence of the same substrates (Thorne *et al*., 2010). In fact, it has been shown that the complete removal of the C- terminal domain while luminescence can still be observed, the N- terminal domain continues to bind the substrates with a lesser affinity and overall bioluminescence is a mere 0.03% to that compared to the non-truncated structure (Zako *et al*., 2003).

The superfamily to which the FLuc belongs utilizes a domain alternation catalytic strategy whereby the first adenylation is conducted via one conformation whilst the second decarboxylation reaction is conducted within an alternative conformation (Gulick *et al.*, 2003; Reger *et al.,* 2008). Following adenylation, there is a 140^{O} rotation of the C-terminal domain which allows the second conformation to be utilized for reaction purposes (Figure 1.3). This feature has been confirmed to take place within FLuc both via biochemical means (Branchini *et al.,* 2000; Branchini *et al*., 2005) and structural methodology (Sundlov *et al.,* 2012).

The active site itself, responsible for the binding of the substrate and the co- factors which are required for bioluminescence, may be elucidated via identification of conservation between the superfamily of adenylating enzymes (Marahiel and Stachelhaus, 1997). Amongst this family, the active site is comprised of 10 conserved regions termed the A1-A10 motifs (Marahiel and Stachelhaus, 1997) and these motifs are important in either or both the adenylation and decarboxylation reactions. The specifics of the residues comprising the FLuc active site have not been conclusively determined however within

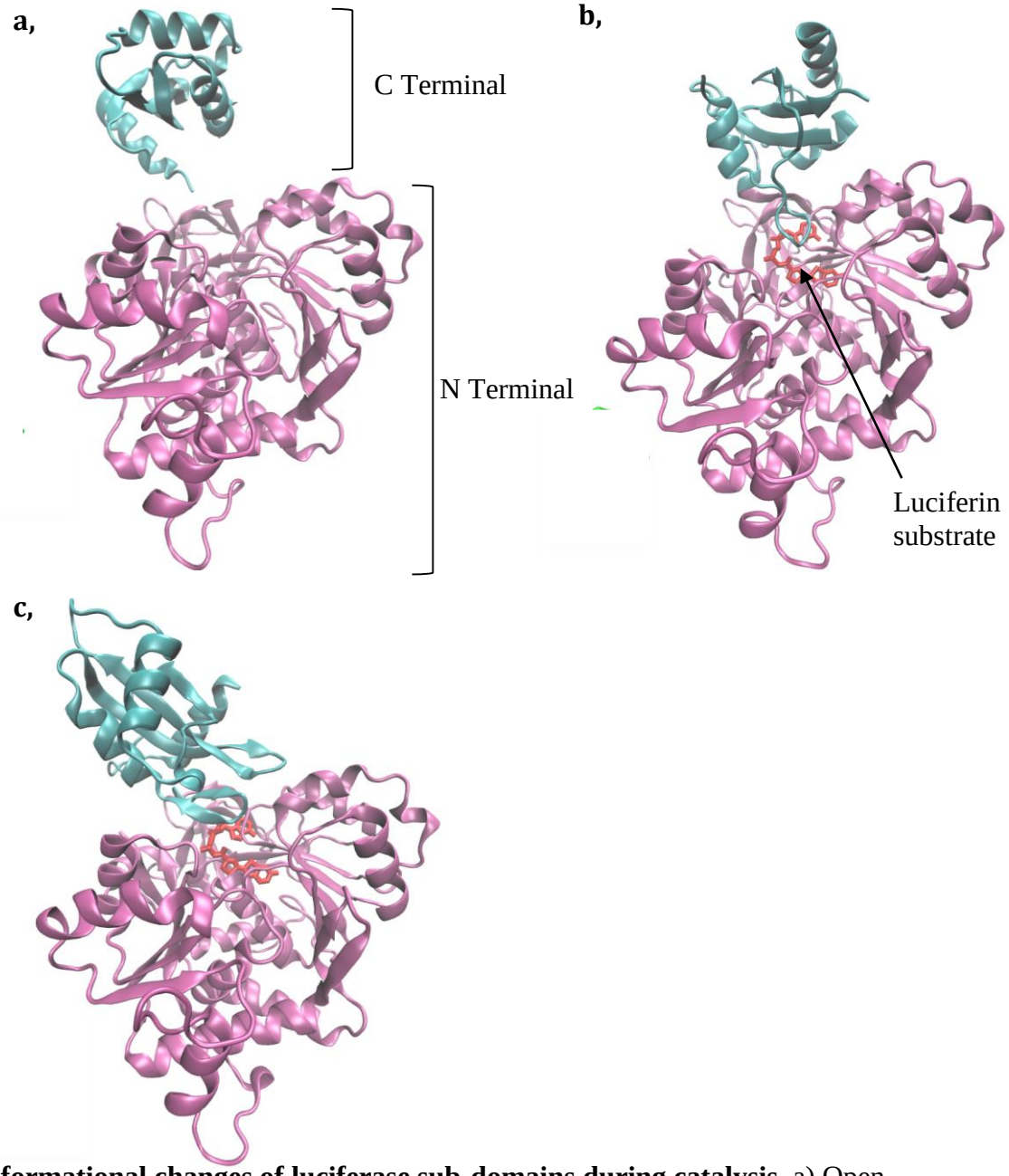
the last decade there has been much progression that has been made within this field which began with the development of a working model of the FLuc active site based on the X-ray structure of the enzyme without bound substrates (Branchini *et al*.,1998). Based on this working model, LH_2 binding studies have identified positions which may be involved in the active site of FLuc (Branchini *et al*., 2003). Site directed mutagenesis was conducted to mutate 15 potential active site residues which were positioned within 5Å from the active site. Characterization of these mutants identified that 12 had altered binding affinities for LH_2. Residues that have been further characterised include Lys 529, Thr 343 and His 245 (Branchini *et al*., 2000). Lys 529, a particularly well conserved residue appears to be critical in determining correct orientation of the substrates since it provides beneficial polar interactions (Branchini *et al* 1998).

Figure 1.2. Secondary Structure of Firefly Luciferase



Secondary structural elements within firefly luciferase (Conti *et al.*, 1996). The overall secondary structural elements are highlighted coloured via subdomains. Arrows are representative of β-strands and are numbered in order whilst circles are representative of the α-helices and are also numbered in order. All secondary structure elements were determined with the database of secondary structure assignments (DSSP).

Figure 1.3. Changes in the structure of luciferase promoted by conformational changes as a result of substrate binding



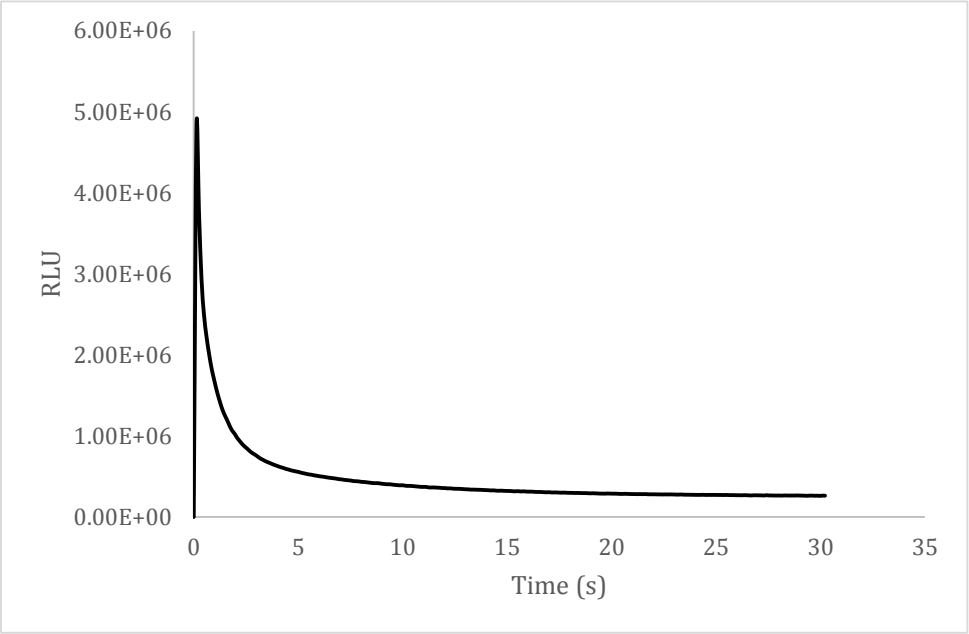
Conformational changes of luciferase sub-domains during catalysis. a) Open conformation b) Adenylate forming conformation c) Second catalytic conformation. On binding of the substrate within the hydrophobic pocket (red) the C- terminal (cyan) moves down on top of the N- terminal (purple) pocket and following adenylation, rotates 140^{o} to undergo further catalysis. Image generated via VMD using PDB files ID: 1LC1, 4G36, 4G37.

1.1.5. Characteristic Kinetics of Firefly Bioluminescence

FLuc exhibits a characteristic profile of light output over time in the presence of the substrates at saturating conditions (DeLuca and McElroy, 1974; Ugarova, 1989) wherein there is an initial flash followed by a steady-state glow (Figure 1.4.). These kinetics arise due to the conformational changes that occur within the enzyme in combination with end product inhibitors of the reaction. Prior to the production of light, there is an initial lag time which lasts for ca. 25ms, which was initially hypothesized to be due to the independent binding of the substrates, ATP and LH_2 after which a ternary complex is achieved and it is this that allows for a rapid rise to the peak intensity (Imax) at 300ms (DeLuca, 1976) . However, a prominent study investigating these characteristics concluded that oxygen addition was a key step to this rapid rise in activity and that the initial lag was a result of two rate limiting conformational changes occurring during the formation of the adenylate to enzyme complex (DeLuca and McElroy, 1974). It is hypothesized that the conformational changes that occur place the enzymatic proton acceptor close to the 4C of the LH_2-AMP in readiness for the abstraction of the proton (Branchini *et al*., 1998) along with the movement of the two domains to enclose the substrates within the active site (as discussed previously see Section 1.1.4.).

The characteristic flash kinetic wherein following the rise to Imax, there is a rapid decay in light output, is primarily due to end product inhibition promoted by oxyluciferin. However inhibition by oxyluciferin does not wholly account for the rapidity of decay observed (Ugarova, 1989). The decay can be considered as a multistep process, due to this initial inhibition by oxyluciferin in combination with dark reaction products of FLuc, the potent inhibitor dehydroluciferyl adenylate (Lemasters and Hackenbrock, 1977). Secondly, bioluminescence, although reduced, continues as a factor of various inactivation pathways in combination with the amount of remaining LH_2 (DeLuca and Mc Elroy 1974; Brovoko *et al.,* 1994).

Figure 1.4 Typical flash kinetic observed during luciferin-luciferase mediated reactions



Characteristic profile of light observed during initiation of the bioluminescent reaction. Bioluminescent activity measured over 30s from wildtype *Ppy* FLuc (WTFLuc) in the presence of the substrates ATP and LH_2, at saturating conditions. Peak activity observed occurring prior to 1s with an Imax of 4.8E+06 RLU with subsequent rapid decay as a result of end product inhibition along with other inactivation pathways.

1.1.6. Further Reactions of Firefly Luciferase

The activity of FLuc, most likely due to its ancestry, has applications not only restricted to the bioluminescence pathway. Due to the heritage of the enzyme, FLuc can act as an acyl-CoA synthetase, and this function appears to be is vital in certain of the side reactions. Indeed, its known activities also are now involved in the LH_2 synthesis pathway, production of the inhibitor dehydroluciferin and may convert long chain fatty acids into fatty acyl-CoA (Nakamura *et al*., 2006; Fontes *et al*., 199; Fraga *et al*., 2005; Oba *et al*., 2003).

LH_2, has two potential enantiomers, the D- and the L-LH_2. The specific chirality has been described to be critical for bioluminescence, since beetle FLucs show high bioluminescence specificity with the D enantiomer whilst the L enantiomer acts as a competitive inhibitor for the reaction (Seliger *et al*., 1961; White *et al*., 1961; McElroy and Seliger, 1962). FLuc can utilize its acyl-CoA synthetase function to convert L-LH_2 to luciferyl-CoA (Nakamura *et al*., 2005). As such, it is considered that both the D- and L-LH_2 substrates inhibit the reaction of its corresponding enantiomer.

Luciferase is involved in the generation of D-LH_2 by slow stereoisomerisation of L-LH_2. Niwa et al (2006) considered that the biosynthetic pathway of D-LH_2 may be formed from L-cysteine and provided evidence suggesting that firefly luciferase mediated the steroisomeration of L- LH_2 into D- LH_2 using CoA. As previously discussed, as a CoA ligase, FLuc may act upon L-LH_2 giving rise to racemic luciferyl-CoA which subsequently could yield both forms of the substrate.

As noted previously (see Section 1.1.5), typically FLucs display a characteristic profile of light emission however there are mechanisms whereby this kinetic may be altered. The production of inhibitor dehydroluciferyl adenylate may be released from FLuc, lifting inhibition via the addition of CoA (Airth *et al*., 1958; Fontes *et al*., 1998). It is considered that FLuc itself may catalyse both the adenylation of the dehydrolucferin and the subsequent conversion into deyhroluciferyl- CoA (L-CoA), removing AMP, whose

adenosine moiety blocks the access of substrates to active site. L-CoA whilst an inhibitor of the reaction, is substantially less potent than the initial LH_2 derivative (Ford *et al*, 1995). This being the case, CoA lends itself to utility within applications requiring sustained light output. D- LH_2 inhibits the luciferyl-CoA synthetase activity of L-LH_2.

The complexity of the reactions in which FLuc may take part can be further broadened to those reactions involved in the synthesis of dinucleoside polyphosphates, able to catalyze the transfer of a nucleotidyl moiety via nucleotidyl-containing intermediates, with release of pyrophosphate. This reaction is considered as part of the pathway involving the formation of dehydroluciferyl adenylate since it appears to act as an intermediate for the synthesis nucleoside polyphosphates (Sillero and Sillero, 2000).

1.1.7. Bioluminescence and Colour Modulation

Conventional bioluminescence of *Ppy* WTFLuc typically emits with a peak emission (λmax) of 550nm-560nm corresponding to yellow-green light and is pH-dependent (McElroy and Seilger, 1961). However alterations in this colour of emission in nature and by protein engineering are common and the mechanisms contributing to such variations are wide ranging.

1.1.7.1. The Bathochromic Shift of Firefly Luciferase

The bathochromic shift is characteristic of WTFLuc in response to denaturation, wherein the light emission from the emitter broadens and becomes red-shifted either due to low pH, high temperature or the presence of heavy metals, denaturants and other various ions. This effect is noted wholly within FLuc and does not extend to other beetle luciferases such as click beetles and railroad worms indicative that the bathochromic shift is a protein mediated event. It is considered that in order for this effect to occur, there must be at least two different conformations of the protein within altered environments of the emitter. Mutagenesis studies identified that binding of ATP may be critical to this effect, highlighting that stabilization made via interaction by the 2' hydroxyl of ATP may favour green light emission (Tisi *et al*., 2002). It is also noted that in the unbound state, there is a

salt bridge set up between residues R437 and E455 (Tisi *et al*., 2002) and denaturation disrupts these interactions and perhaps alter the ability of ATP to bind correctly (Tisi *et al*., 2002).

1.1.7.2. Bioluminescence Emission Colour Mechanisms

The colour of bioluminescence may be modulated by FLucs, with wide shifts from λmax 560nm to 610nm for *Ppy* FLuc and the mechanism for this has been heavily studied.

The colour-shifting mechanism can be categorized into one of two separate types of shift, shifts caused by specific or non-specific effects on the emitting species. Specific effects comprise those spectral changes whereby there is a change in the spectral full width at half maximum (FWHM), whilst nonspecific effects are those whereby there is a no change in the spectral FWHM however there is change in the peak emission (Ugarova and Brovko, 2002). These effects can be invoked by different mutations in the structure of FLuc.

Specific effects whereby the spectra narrows or broadens are due to the degree of freedom of the oxyluciferin emitter and provides information on the variable forms that the emitter may take. The spectral shape has a predominant peak corresponding to the most common form of the emitter state of the oxyluciferin however surrounding this peak there is an asymmetrical distribution attributing to the other various potential emitting species. Bandwidths that are narrower have greater specificity, rigidity of the emitter and tend to have higher stabilization within the active site. Specific effects are due to direct stabilization of LO* emitter forms via acid-base interactions, H-bonds or electrostatic interactions. Changes as a result of nonspecific effects are considered to be due to solvent-like global shifts of all emitter forms (Ugarova, 2008). Non-specific effects have been described in terms of orientation polarizability of the emitter(s) (Ugarova and Brovko, 2002).

Specific effects can also be sub-categorised. Branchini *et al* (2004) have proposed a model based on the observations made by White *et al* (1980) in which it is considered

that there are two forms of the emitter, the keto and enol forms of the 4-hydroxyl of LO*. However, various other studies have concluded that this theory is unsupported (Branchini *et al.*, 2002; Nakatani *et al.,* 2007). 5,5-dimethylluciferin was utilized which blocks enolization of the emitter and this led to red or green emission. Utilising a 2-(4-toluidino) naphthalene-6-sulfonate probe it was identified that the LH_2 binding site is polar when emitting red shifted bioluminescence when compared to green-yellow emissions. This suggests that the keto form of the excited oxyluciferin is the true emitter whilst the colour of emission is modulated by the 6'-hydroxy group of oxyluciferin interacting with basic residues (specific effect) and the polarity of the active site (non-specific). Branchini proposed and alternative hypothesis termed the resonance-based charge delocalisation mechanism to explain differences in colour and this has been supported by quantum mechanical calculations and recent chemical evidence (Jathoul *et al.*, 2014).

1.1.7.3. Modulation of Colour Utilizing Luciferin Analogues

Bioluminescence may be modulated simply via the utilization of analogues of the LH_2 substrate (Sun *et al.,* 2012). Various LH_2 analogues have been synthesized which have inherently altered emission. Jathoul *et al*. (2014), synthesized a dual color, far-red to near-infrared analogue by increased conjugation within the LH_2. Similar to native LH_2 it exhibits pH dependent spectra and emits differently dependent on the mutant FLuc enzymes with a maxima of up to 706nm. Aminoluciferin (White et al., 1966) is another analogue with which FLucs typically emit red shifted light with a higher catalytic efficiency (k_{cat}/K_M) than native LH_2.

.

1.1.8. Current Protein Engineering Strategies Utilized for Improved Characteristics of Firefly Luciferase

WT *Ppy* FLuc is highly thermolabile (Tisi *et al.,* 2002). In addition to this, WT has a half-life of 3-12 hours (Auld *et al*., 2009), which is advantageous for investigating dynamic responses in gene reporter systems, is a disadvantage for *in vitro* diagnostics since the protein activity is heavily dependent on the environment in which it is used. This

instability has impeded its utility within academia and industry, in particular, instability of this enzyme towards environmental factors, such as temperature and pH. Stabilizing compounds can in part ease these issues, but protein engineering strategies have been extensively employed to overcome these fundamental issues with FLuc. Strategies that have been employed through mutagenesis efforts range from substitution mutagenesis to chimerisation of different enzymes and have led to the development of a bank of FLucs with altered emission peaks, enhanced activities and greater stability towards temperature and also common inhibitory agents, such as pH, the presence of solvent, ions and other denaturants (White *et al*., 1996; Tisi *et al*., 2002; Law *et al*., 2002; Jathoul *et al*., 2012; Branchini *et al*., 2014).

Substitution mutation strategies have given rise to mutants with wide ranging characteristics, e.g. kinetics, spectra and enhanced stability by screening of mutant FLucs expressed within bacterial colonies aimed to identify desirable mutants. Novel mutants are not limited to changes close to the active site of the protein, but involve a myriad of mutations occurring throughout the protein from the core to the outer surface. For example, changes to the outer surface can cause non-specific effects on the emitter while leaving the catalytic activity unperturbed (Ugarova and Brovoko., 2002; Law *et al*., 2006). In addition to this, the Luc genes have been manipulated via codon optimisation for increased expression within bacteria or human expression systems or cells (Sherf and Wood, 1994; Davies *et al*., 2010).

Thermostabilising mutations have revolutionized the applications of FLuc. Historically, thermostability refers to thermodynamic stability which reflects the initial and final states of a protein and is commonly expressed as Gibbs free energy of unfolding. However the term is now commonly used with reference resistance to thermal inactivation of FLuc, meaning stability to high temperatures is conferred, but this may be heavily influenced by both kinetic stability, which is a measure of the activation of free energy and thermodynamic stability, is a measure of unfolding of the protein at equilibrium either due to heating or denaturant inclusion.

Improvements made to the stability of enzyme, as is often seen amongst other proteins, typically involve a trade-off between thermostability and activity (White *et al.*, 1996; Jathoul *et al.*, 2012) that by improvement of thermostability, the rigidity imposed upon the protein can reduce the activity or the catalytic efficiency. This trade-off is normally compensated by running the reaction at higher temperatures. The impact of thermostabilising mutations can also alter the landscape of protein folding and prevent aggregation.

Ultra-Glo, as patented by Promega (Promega Corp., Madison, WI, USA), is suggested to contain up to 27 mutations within *Photuris pennsylvanica luc* (Amit Jathoul, PhD Thesis, University of Cambridge, 2008) and is considered as one of the most thermostable FLucs engineered to date (Hall *et al*., 1999). As a result of directed evolution, whereby selection was based on a number of criteria including overall bioluminescence, thermostability and quantum yield, the mutagenesis methodologies employed to give rise to extensive libraries on which selection could be made included site directed mutagenesis and DNA shuffling. Throughout characterization a mutant was isolated able to tolerate 50^{o}C for a total of 5 days whilst kinetically the mutations had had no impact on the catalytic mechanism and as such, additional investigation was conducted to isolate mutants whereby selection was based on product inhibition reduction through which Ulta-Glo was discovered.

Characteristically, Ultra-Glo is highly stable, able to withstand 65^{o}C for up to 5 hours, has high affinity for the LH_2 substrate and over time has a higher overall light yield in comparison to its predecessors (Hall *et al*., 1999). Whilst Ultra-Glo has a predominantly lower flash kinetic, the decay of signal is also reduced showing that Ultra-Glo is more so resistant to product inhibition (Hall *et al*., 1999).

1.1.8.1. The Development of Thermostable and pH Tolerant *Ppy*, x11Fluc

Combinations of mutations have previously been shown to amplify the characteristics caused by individual substitutions. E354K and E354R were both implicated in increasing thermostability (White *et al*., 1996) considered to be due to a loss of the negative charge

at this position. These mutations may also have a role in stabilization of the active site since E354I caused a red shift in the peak emission (White *et al*., 1996). Further work then identified thermostability mutations D357Y and D357F (White *et al*., 2002). Combining these gave rise to mutants E354R/D357Y and E354I/D357Y. Despite similarities in their makeup, E354R/D357Y termed x2FLuc (Table 1.1.) resisted the bathochromic shift at low pH whilst E354I/D357Y was red-shifted with wider bandwidth throughout all pH values.

Further thermostabilised mutants have been developed with the hypothesis that reduction in the burial of hydrophilic residues that there would be an increase in the packing of the enzyme (Prebble *et al.*, 2001) and through combinations of residues identified as either hydrophobic and hydrophilic, x4FLuc was developed (Tisi *et al.*, 2002) containing mutations T214C, I232A, F295L E354K, which enhanced thermostability and resistance to bathochromic shift, albeit with lower specificity activity and k_{cat} and an increase in the K_M for the LH_2 substrate compared to WTFLuc.

In 2002, Law *et al* utilized positions previously identified to confer thermostability when mutated to alanines (Tisi *et al.*, 2001) and substituted these positions to other residues. Specifically, the residues selected were surface exposed and initially hydrophobic which throughout the study were mutated to hydrophilic residues. It is suggested that mutations of this fashion would increase the surface polarity. One mutant was derived from the study and termed x5 containing mutations F14R, L35Q, V182K, I232K and F465R (Table 1.1). x5 was more resistant to thermal inactivation at higher temperatures than previous mutants, without an increase the K_M for the substrate or lowered specific activity compared to WTFLuc derived from *Ppy*.

Further work led to the development of x12FLuc (Table 1.1), an enzyme constructed by combination of the x2, x4, and x5 mutations and with 1 additional mutation. This mutant was not only resistant to thermal inactivation (retaining 85-90% of initial activity after 1 hour at 40ºC whilst x2 and x4 were reduced to 75-80% and x5 reduced to 20%), but was also more tolerant to alterations in activity with fluctuations in pH, with 80% of

maximum activity retained by the enzyme over a pH range of 6.6-8.6 (Jathoul *et al.*, 2012). It must be noted that all these mutants were more stable than the WT which rapidly deactivates within 10 minutes at the same temperature.

Following on from this work, two revertants of x12FLuc were made: T214C and F295L (Jathoul *et al.*, 2012). Whilst the T214C revertant showed little dissimilarity to the template x12, F295L had a lower K_M for LH_2, a higher k_{cat} and a resulting higher k_{cat}/ K_M ratio. The enzyme was termed x11 firefly luciferase (x11Fluc) (Table 1.1) and in summary contains 11 mutations, 9 surface and 2 internal mutations exhibiting K_M of 3.7±0µM for LH_2, compared to 18±1µM for the WT. The specific activity in particular was two-fold higher than x12FLuc whilst still retaining similar activities at both 40 ºC and 50 ºC. An enzyme such as this has wide ranging implications since it is more beneficial compared to previous mutants in terms of sensitivity at elevated temperatures, for example as an *in vivo* reporter gene for studies in mice.

Table 1.1. Summary of mutants involved in the development of x11FLuc (modified from Jathoul *et al*., 2012)

Mutation	**Mutants**				
	x2	**x4**	**x5**	**x11**	**x12**
F14R			+	+	+
L35Q			+	+	+
A105V				+	+
V182K			+	+	+
T214A		+			
T214C				+	+
I232A		+			
I232K			+	+	+
D234G				+	+
F295L		+			+
E354R	+			+	+
E354K		+			
D357Y	+			+	+
S420T				+	+
F465R			+	+	+

Summary of Mutants involved in the development of x11FLuc. Mutations present within each of the 4 mutants, x2FLuc, x4FLuc, x5FLuc, x12FLuc which were precursors to x11FLuc are presented for comparison to x11FLuc.

1.1.9. Applications for Bioluminescence

Today, light emitting systems are used as indispensable analytical tools in various fields of science, technology and medicine from GFPs to phycobiliproteins to the bioluminescent proteins all of which have their own advantage and disadvantages. GFPs are heavily utilised as useful marker proteins due to the intrinsic fluorescence (Tsien, 1998) whilst phycobiliproteins, proteins covalently linked to a fluorescence group, are employed for use in flow cytometery amongst others (Telford *et al*., 2001). Currently, bioluminescence is key in a variety of fields and improvements to this system has been actively pursued. This popularity of luminescence has become more prominent since unlike fluorescence it requires no energy for excitation allowing for imaging of deep tissues (DeLuca and McElroy, 1978; Jathoul *et al.*, 2014), it is non-toxic in mammalian cells, it can deliver a high signal to background ratio providing a highly sensitive assay. Such applications include but are not limited to; High Through-Put Screening (HTS) assays, pyrosequencing, *in vivo* imaging and as a reporter for gene expression (Ronaghi, 2001; Branchini *et al*, 2009; Taurianen *et al*., 2000). The degree to which luminescence systems have been implicated in new technologies may be demonstrated by the circa. 420 bioluminescent assays listed within the PubChem database, approximately, 21% of the total assays listed.

Red shifted variants of FLuc tend to have lower quantum yields of bioluminescence, however, use of bioluminescence for bio imaging is becoming an important tool due its high sensitivity, which is part due to high signal to noise ratios *in vivo*. As discussed, the limitation of FLuc bioluminescence for *in vivo* imaging due to the restriction imposed by the D-LH_2 substrate whereby haemoglobin in mammal tissues absorbs visible wavelengths below 600nm and this reduces the effectiveness of the native system as *in vivo* imaging tool in deep tissues. To overcome the restricted wavelengths of emission of native LH_2, novel LH_2 analogues have been synthesized to incorporate extra conjugation with the resulting molecule giving emission shifted into the near infra-red (Iwano *et al*., 2013; Jathoul *et al.*, 2014).

1.1.9.1. Firefly Luciferase as ATP Sensors

There are many differing methods in the quantification of ATP such as QUEEN (Yaginuma *et al*., 2014), a genetically-encoded ratiometric fluorescent ATP indicator, however to date, FLuc bioluminescence remains to be the most sensitive, reliable and most practical technique (Branchini *et al.*, 2015). FLuc is sold commercially as a sensitive and accurate assay for ATP and it is able to detect ATP down to femtomolar (10^{-15} M) concentrations. This ATP-sensing behaviour is important in a plethora of technologies.

A common use for luciferase-based ATP quantification within the commercial sector is in continual hygiene monitoring practises within a number of industries, ranging from restaurants to hospitals to those companies involved in the manufacture of food. ATP quantification has been adopted by these industries in the measurement of bacterial contamination since ATP is the universal energy carrier of cells and indicates the presence of microbes. To this end, all hygiene monitoring tests based on this principle inevitably also test not only for bacteria, but also other contaminating substances such as the presence of blood or food matter.

1.2. Dogma and New Directions in Firefly Luciferase Protein Engineering

1.2.1. The Central Dogma

Within nature there is an abundance of proteins which are diverse in their sizes, shapes and functionality. Principally these proteins are comprised of 20 natural amino acids, and the central dogma dictates that DNA is converted to mRNA which encodes the exact sequence of amino acids that comprise the polypeptide chain, which ultimately dictate the structure, folding and subsequent function of the protein. This inherent structure, folding, function relationship allows the modification of proteins, which not only has enabled the elucidation of underlying mechanisms to determine protein function, it has led to the development of the field of protein engineering.

1.2.2. Protein engineering

Protein engineering, has been crucial for the construction of novel variants of many types of enzymes in biotechnology, lending itself to the redesign and improvement of existing catalysts in research and industry (Brannigan and Wilkinson, 2002; Baldwin et al., 2008). Additionally, it has acted as a tool to understand the mechanisms through which enzymes function (Brannigan and Wilkinson, 2002) and is quoted to be “an extremely powerful tool” within fields concerned with the link between sequence, structure, folding and function (Price and Nagai, 2002).

1.2.3. Strategies for Protein Design and Engineering

It is commonly understood that there are two major methods by which proteins can be adapted to generate new variants: rational protein design or directed evolution. Both of these methods have their limitations as well as advantages (Lutz and Patrick, 2004), for example rational site-directed mutagenesis (SDM) requires prior structural knowledge of the protein while directed evolution requires the construction of extensive libraries containing many different variants of the protein being studied. Despite this limitation,

following the generation of a library of variants, screening those proteins with desired characteristics to identify beneficial mutations makes directed evolution a very powerful tool in the adaptation of proteins. Conventional methodologies that have been employed in generating such mutants include; random mutagenesis (Koksharov and Ugarova, 2008), DNA shuffling (Maguire *et al.*, 2010) and to a lesser extent circular permutation (Osuna *et al*., 2002), a technique involving the intramolecular relocation of the N- and C-termini of a protein and yet to date mutagenesis involving insertions and to a greater extent deletions are rarely employed.

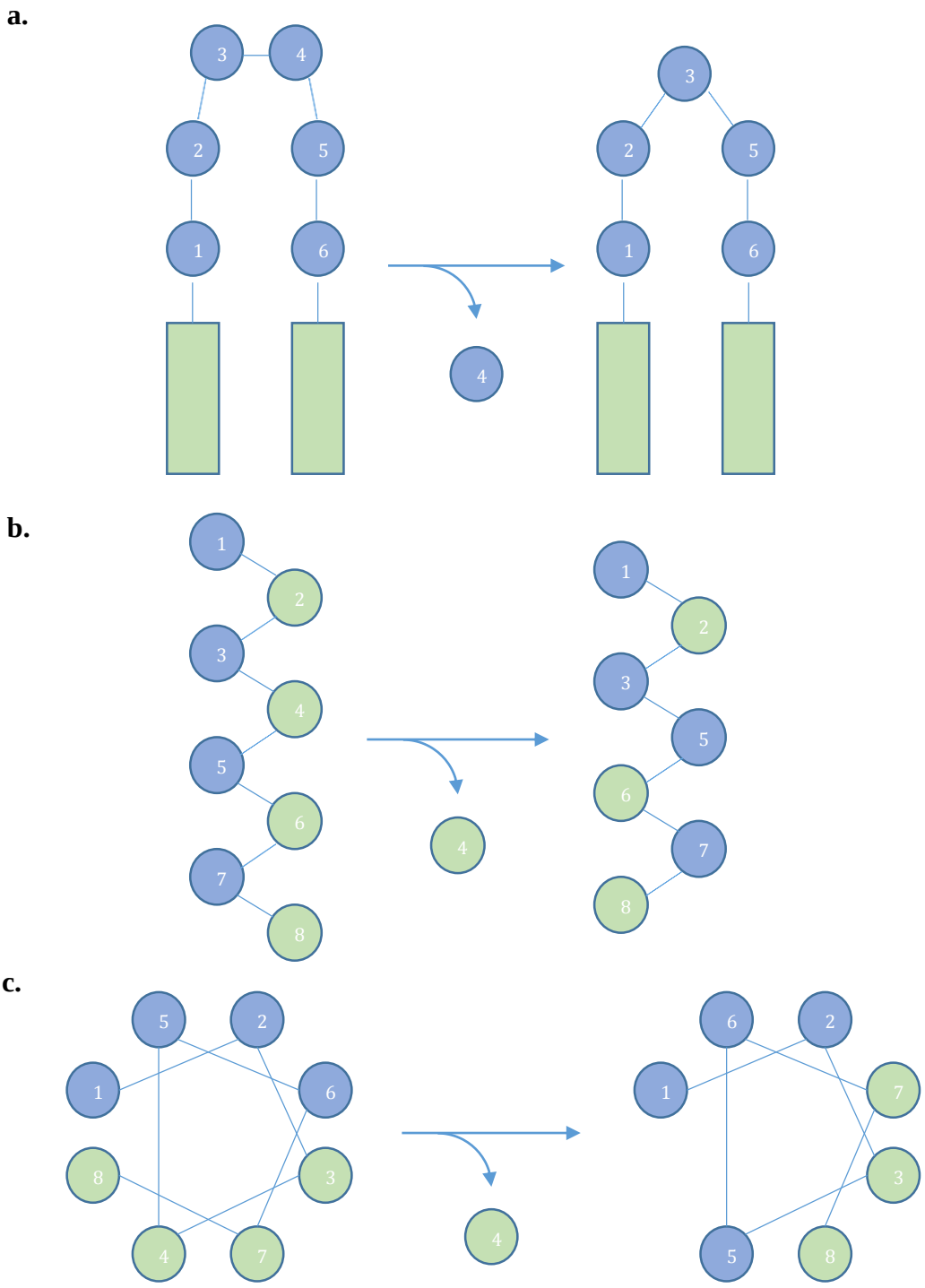
In nature, the insertion and deletion (InDel) mutations are some of the most common stressing that this type of mutation may be an important an evolutionary mechanism in nature and therefore may offer an alternative to the current tools utilized within the field of protein engineering (Taylor *et al*., 2004). Substitution mutations, which are most commonly utilized, affect only the side chain of the amino acid, allowing for diverse changes at the target site, however, the backbone of the protein remains unperturbed. Insertions and deletions on the other hand, inherently alter the backbone of the protein due to the addition or removal of amino acid residues and this is considered to be far more destabilsing to a protein. In addition to this, whilst the structure of a protein as a result of substitutions can largely be predicted by molecular visualization and analyses, changes as a result of the InDel mutations are difficult to predict with current software. However, in spite of these difficulties, whilst substitution mutations restrict the conformational space that a protein may assume, insertion and deletions allow for greater sampling and provide the potential for the development of novel variants.

Proteins are thought to be less tolerant to deletions in comparison to the insertions since insertions can increase the conformational space, however, deletions may reduce it. However, within nature deletions occur up to three times more often than insertions (Zhang and Gerstein, 2003) and in light of this argument perhaps deletions are not so deleterious to function as once considered. This frequency for which deletions occur may be understood via understanding of the repair mechanism of DNA. Commonly a piece of ssDNA will loop out as a result of a change to the base pair sequence. The mismatch that

arises between the ssDNA and the complementary strand will tend to target the loop for cleavage promoting the deletion of base pairs.

InDels of amino acids in certain locations of proteins are significantly well tolerated compared to other locations. At the level of DNA, firstly, InDels of nucleotides may promote a frame shift of the entire gene and this is far more deleterious to the functionality of the protein, however those InDels occurring in multiples of three base pairs, bypass such frame shifts. In addition to this, it is known that certain elements within the protein are more so tolerant to the registry shifts that occur as a result of such mutagenic methods. Loops are such an example owing to larger conformational flexibility in comparison to other structural elements whereby registry shifts are thought to impact the protein to a greater extent, as illustrated in Figure 1.5. Within secondary structural elements, where a single amino acid deletion occurs within a β-strand, the side chains shift to opposite faces of the strand, promoting the loss of previous interactions and the introduction of new interactions. Another registry shift that may occur is as a result of a single amino acid deletions within α-helices whereby as a result of a deletion there is a rotation in the position of the side chains. As discussed, the ubiquitous nature of InDels within nature, highlights the importance of InDels as part of evolution. For example, InDels within human antibodies against the HIV envelope protein are critical for the variety of epitope recognition sequences within its hypervariable regions (Kepler *et al*., 2014). Taken together, despite the dogma that amino acid deletions likely compromise the structural integrity of proteins, the potential novelty that may be imparted by deletions cannot be overlooked. Whilst substitution methodologies are heavily employed to develop new proteins, investigation into the utility of the InDels must be examined.

Figure 1.5. Registry shifts occurring as a result of a single amino acid deletions



The effect of single amino acid deletions within structural elements. a. Deletion within a loop cause a shortening of the loop. **b.** If there is a deletion within a β strand, the side chains shifts to the opposite face of the strand. **c.** Deletions within an α-helix cause a rotation in the position of the side chains.

1.2.4. Current Protein Engineering Incorporating Deletion Mutations

In recent years, despite the dogma surrounding functionality of proteins following deletion mutations, new studies are confirming the utility of these mutagenic methods and highlight the importance of deletions. Indeed, proteins can benefit even when the deletion is within secondary structures. Simm *et al*. (2007) has encouraged this new outlook describing not only toleration of deletions within loops and secondary structures of TEM-1 ß-lactamase, but an enhanced activity towards a ceftazidime substrate compared to the WT, displaying a 64-fold increase in activity of a deletion in P174. Similarly, deletions of six amino acids promoting a closer compacted state within Staphylococcal nuclease increases the catalytic activity and stability compared to the WT (Baldisseri *et al.,* 1991).

Directed evolution in combination with random single amino acid deletion sampling within eGFP has shown that certain deletions are tolerated throughout the protein, suggesting the beta-barrel fold remains intact despite deletions throughout it (Arpino *et al.,* 2014). This study confirmed that loops, helical elements and the termini of strands were more tolerant to single amino acid deletion in comparison to other structures. In addition to this, a variant termed G4 displayed higher fluorescence due to a single amino acid deletion within α-helix. Structural analysis showed that a helical registry shift had resulted in a new polar interaction network, possibly stabilising a *cis* proline peptide bond, indicative that in some cases registry shifts are not detrimental to protein function.

Incorporating extra such diversity into libraries via InDel approaches increases the potential for other strategies such as directed evolution whereby the power of the strategy is wholly reliant on the diversity of libraries to be screened. Shortle and Sondek (1995) stated, “although their effects are difficult to anticipate, insertions and deletions provide important tools for altering protein structures in directions not achievable by substitutions alone” (Shortle and Sondek 1995).

1.2.5. Deletions of Amino Acids within Firefly Luciferase

As is the case for many proteins, whilst there are a plethora of FLuc mutants derived from substitution mutagenesis as has been previously discussed, there are no reported systematic studies which concern the FLucs with regards to the sampling of deletions.

There are two main methods for the incorporation of single amino acid deletions within a protein, these include oligonucleotide based methods or a transposon based technique (Lutz and Patrick, 2004; Neylon, 2004).

In the 1987, Gould *et al* conducted a study whose primary interest involved the movement of proteins throughout the cell and in the elucidation of peptide targeting signals in the transport of proteins into the peroxisomes. Since FLuc is naturally localized within peroxisomes of cells, the study utilized FLuc as a model for their study. By removal of up to 20 amino acids from the terminals of the protein, this study concluded that the signal peptide was no more than 12 amino acids in length and existed at the carboxy-terminus of the FLuc. Despite the generation of deletion mutants, due to the focus of this study centering on the transport system, no characterization of mutant proteins was considered (Gould *et al*., 1987). Other studies involving deletions within FLuc include the cumulative addition of deletion of amino acids MRSAMSGLH from the C- terminus of FLuc (Sala Newby and Campbell, 1994). Removal of up to seven of the C-terminal amino acids did not results in a complete loss of function, however, there was an increasing loss in activity when amino acids 8-12 were removed. Replacement of 539-550 and 543-550 resulted in FLucs that displayed 22% and 35% of activity respectively.

Further studies have investigated deletions within the N- terminal of FLuc (Sung and Kang, 1998; Wang *et al*, 2002). Deletion of amino acid residues 3 to 10 within the N terminal are noted to reduce the bioluminescent activity of the FLuc to less than 1% displayed by the WTFLuc (Sung and Kang, 1998). Additional investigations studying the deletion of the first six amino acid residues demonstrated a 29% loss in the

bioluminescent activity whilst deletion of an additional amino acid residue at position 7 gave less than 0.5% of the original activity.

1.3. Aims and Objectives

Owing to the importance of FLuc within academia, medicine and technology, development of 'toolboxes' of FLuc variants is highly desirable, for example, mutants with altered characteristics such as shifted spectral properties, enhanced stability in various assay conditions. Over the years, FLuc has been heavily engineered, however, to date, all existing mutants of FLuc were generated via SDM. The use of deletions as a method of improving FLuc appears to have been completely overlooked.

The overarching aim of this project, is to employ single amino acid deletions within x11Fluc in order to address two questions. Firstly, if FLuc can tolerate single amino acid deletions and to what extent, and secondly, to identify those mutations which result in desirable characteristics such as increases brightness, altered colour and altered kinetics. Furthermore, screening methodologies will be optimised within this study with the aim to identify desirable characteristics from a large library of mutants, without the necessity of protein purification and for detailed characterisation.

Chapter 2

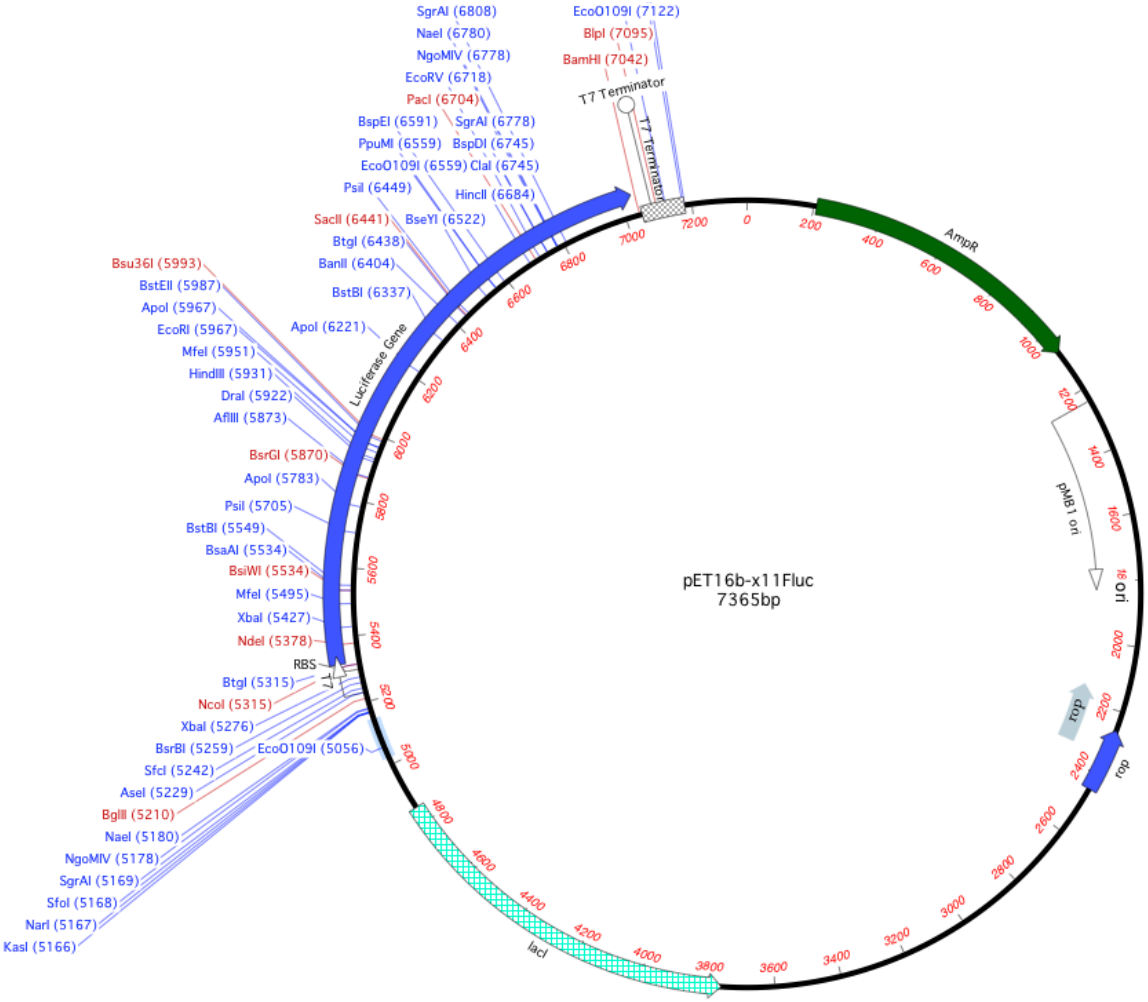
Materials and Methods

2.1. Materials

2.1.1. Chemicals

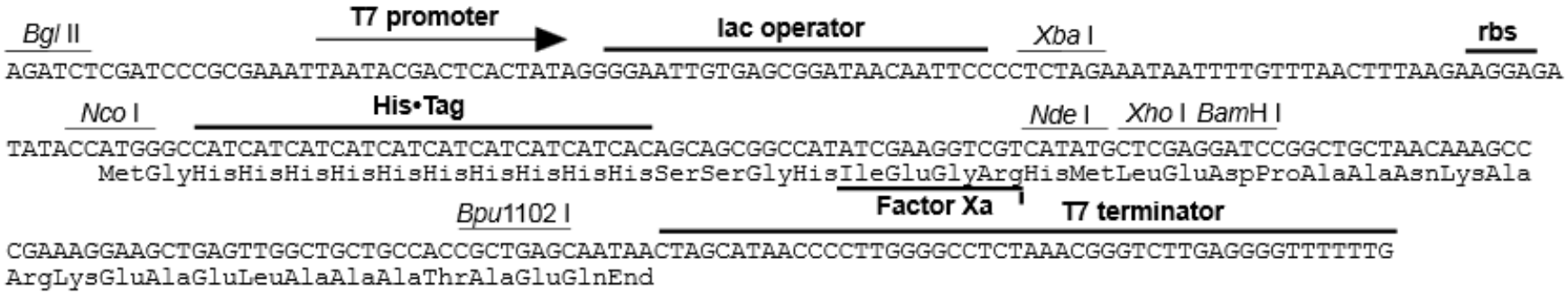
Deionised or MilliQTM (MQ) water were used throughout. Stock solutions of sterile ampicillin (Melford, Suffolk) or carbenicillin (Melford, Suffolk, UK) were made at 100 mg/ml and filter sterilized with a 0.22 μm filter unit (Thermo Fisher Scientific, MA, USA) and were stored at -20^{o}C. Stock solutions of sterile 1M isopropyl β-D-thiogalactopyranoside (IPTG) (Melford, Suffolk, UK) were filter sterilized and stored at -20 ^{o}C. Stock solutions of 31.4 mM D-LH_2 (Europa Bioproducts, Ely, Cambridgeshire, UK) and 100mM ATP (Roche Diagnostics, IN, USA) (pH 7.8) were stored at -20^{o}C prior to use. Both D-LH_2 and ATP were diluted in 'TEM' buffer (100mM Tris-acetate, 2mM ethylenediaminetetraacetic acid (EDTA) and 10mM magnesium sulphate ($MgSO_4$)). The pH of TEM buffer was adjusted with NaOH or acetic acid and stored as a 10x stock at room temperature (RT). Agarose (Melford, Suffolk, UK) was of molecular biology grade and prepared by boiling 1.0% (w/v) in TAE (Tris Acetate EDTA) (Sigma-Aldrich, St. Louis, MO, USA). For the purposes of colony screening, LH_2 and ATP in TEM buffer was diluted in 0.1 M sodium citrate (pH 5.0) to various concentrations.

Figure 2.1. Plasmid Map of pET16b-x11FLuc



Plasmid map of pET16bx11Fluc. AmpR encodes β lactamase; Luc encodes for the *P. pyralis* x11 Fluc luciferase gene; T7 is a promoter that produces high levels of transcription in the *E.coli* BL21 background; lacI encodes for the repression of the T7 promoter. Image generated using MacVector from sequenced pET16b-x11FLuc.

Figure 2.2. Detailed map of pET16b cloning/ expression region



Cloning expression region of pET16b of the coding strand transcribed by T7 RNA polymerase as modified from Novagen pET16b manual (EMD Millipore Corporation, Darmstadt, Germany) highlighting the position and nucleotide sequences of the T7 promoter, lac operator, rbs, the N- terminal 10x His tag and T7 terminator. Position and identity of restriction sites included within this region are *Bgl III, XbaI, NcoI, NdeI, XhoI, BamHI, Bpu1102 I.* Figure adapted from Novagen (2011).

2.1.2. Bacterial Cell Strains and Plasmids

Plasmids containing the relevant FLuc genes of interest were sourced; the two pET16b plasmids containing the WT and x11Fluc genes were kindly donated by Amit Jathoul (University of Cardiff, UK) (Figure 2.1.). pDEST17 containing the x2 insert was sourced from Erica Law. Both plasmids encode ampicillin resistance as a selection marker, with gene expression under the control of IPTG (Figure 2.2.). NEB 5-alpha Chemical Competent *E. coli* (New England BioLabs Inc. (NEB), Ipswich), of genotype fhuA2Δ(argF-lacZ)U169 phoA glnV44 F80? (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17. *E. coli* BL21(DE3) (pLysS) strain of genotype F–, *omp*T, *hsd*S_B (r_B–, m_B–), *dcm*, *gal*, λ(DE3), pLysS, Cm^r were obtained from Promega (Madison, WI, USA).

2.1.3. Bacterial growth media

Lysogeny Broth (Melford, Suffolk, UK) was prepared by dissolving 20g/l of distilled water. LB Agar (Melford, Suffolk, UK) was prepared by dissolving 35g/l of distilled water. SOC prepared comprising Super Optimal Broth (10ml, SOB, Melford, Suffolk, UK) supplemented with filter sterilized 20mM Glucose (Melford, Suffolk, UK). All media were sterilized by autoclave at 121^{o}C. Bacterial growth media were supplemented with 100µg/ml carbenicillin.

2.1.4. Molecular Reagents

Molecular biology reagents were obtained from Roche Diagnostics (USA), NEB (Ipswich) and Sigma Aldrich (St. Louis, MO, USA). Restriction enzymes and complementary buffers were obtained from NEB (Ipswich). DNA Ladder (1 Kb, NEB) was prepared to a final concentration of 100 ng/µl in deionised water and 1x loading dye (6x, New England Biolabs, MA, USA). A total of 5µl was loaded onto gels. Protein molecular weight markers were obtained from NEB and used according to the manufacturer's protocol. Stock 10mM dNTPs solutions were prepared by premixing 4x individual 100mM NTP solution and diluted in water.

2.1.5 Oligonucleotide primers

Stock solutions of all primers were prepared as a 100mM following supply as a lyophilised stock from Sigma and were reconstituted with ultra-high purity water. Short-term stock solutions of all primers were diluted to 10µM and as with longer term primers, all were stored at -20^{o}Cuntil required.

2.1.5.1. Sequencing

Primers (Table 2.1.), were utilized for use in sequencing throughout the whole study.

2.1.5.2. Mutagenesis

Primers (Table 2.2.), were utilized for incorporating the deletions within the *luc* gene within x11Fluc pET16b via site directed mutagenesis.

Table 2.1. Sequence of primers used for sequencing

Primer Notation	**Primer Sequence (5'-3')**
pEXP-F	5'–TGCTCACATGTGCGTAGAGG-3'
DDJ013	5'CAGGGTTATTGTCTCATGAGC-3'
T7 Forward	5'-TAATACGACTCACTATAGGG-3'
T7 Reverse	5'- GCTAGTTATTGCTCAGCGG-3'

Sequence of primers used for sequencing. Primers are listed including the name of the primer and its corresponding sequence.

Table 2.2. Primers Designed to Incorporate a Single Amino Acid Deletion at Specified Sites.

Amino	**Forward**	**Forward Primer Sequence**	**Reverse**	**Reverse Primer Sequence**
M1	FOR1	gaagacgccaaaaacataaagaaagg	REV1	atgacgaccttcgatatggcc
E2	FOR2	gacgccaaaaacataaagaaaggcc	REV2	catatgacgaccttcgatatggcc
D3	FOR3	gccaaaaacataaagaaaggcccg	REV3	ttccatatgacgaccttcgatatgg
A4	FOR4	tttctttatgtttttggcgtcttccata	REV4	aaaaacataaagaaaggcccggc
K5	FOR5	gtcttccatatgacgaccttcgatatg	REV5	aacataaagaaaggcccggcac
N6	FOR6	ggctcttccatatgacgacc	REV6	ataaagaaaggcccggcac
I7	FOR7	tttggcgtcttccatatgacgac	REV7	aagaaaggcccggcac
K8	FOR8	gtttttggcgtcttccatatgacga	REV8	aaaggcccggcacca
K9	FOR9	tatgtttttggcgtcttccatatgacg	REV9	ggcccggcaccac
G10	FOR10	ctttatgtttttggctgcttccatat	REV10	ccggcaccacgctatcc
L172	FOR172	cctcccggttttaatgaatacgat	REV172	atgagatgtgacgaacgtgtacatc
P173	FOR173	cccggttttaatgaatacgatttt	REV173	tagatgagatgtgacgaacgtgtac
P174	FOR174	ggttttaatgaatacgattttaaaccagaaa	REV174	aggtagatgagatgtgacgaacg
G175	FOR175	tttaatgaatacgattttaaaccagaaagc	REV175	gggaggtagatgagatgtgacg
F176	FOR176	aatgaatacgattttaaaccagaaagct	REV176	accgggaggtagatgagatgtg
N177	FOR177	gaatacgattttaaaccagaaagctttg	REV177	accgggaggtagatgagatgtg
E178	FOR178	tacgattttaaaccagaaagctttgat	REV178	attaaaaccgggaggtagatgagatg
Y179	FOR179	gattttaaaccagaaagctttgatcg	REV179	ttcattaaaaccgggaggtagatgag
D180	FOR180	tttaaaccagaaagctttgatcgtg	REV180	gtattcattaaaaccgggaggtagatg
F181	FOR181	aaaccagaaagctttgatcgtga	REV181	atcgtattcattaaaaccgggaggtag
K182	FOR182	ccagaaagctttgatcgtgac	REV182	aaaatcgtattcattaaaaccgggag
P183	FOR183	gaaagctttgatcgtgacaaaaca	REV183	Tttaaaatcgtattcattaaaaccggg
E184	FOR184	agctttgatcgtgacaaaacaatt	REV184	Tggtttaaaatcgtattcattaaaaccg
S185	FOR185	tttgatcgtgacaaaacaattgc	REV185	ttctggtttaaaatcgtattcattaaaacc

F186	FOR186	gatcgtgacaaaacaattgcac	REV186	Gctttctggtttaaaatcgtattcattaaa
D187	FOR187	cgtgacaaaacaattgcactg	REV187	aaagctttctggtttaaaatcgtattca
R188	FOR188	gacaaaacaattgcactgataatga	REV188	atcaaagctttctggtttaaaatcgtattc
D189	FOR189	aaaacaattgcactgataatgaattcc	REV189	acgatcaaagctttctggtttaaaatc
K190	FOR190	acaattgcactgataatgaattcctc	REV190	gtcacgatcaaagctttctggtttaa
T191	FOR191	attgcactgataatgaattcctctg	REV191	tttgtcacgatcaaagctttctg
T352	FOR352	ccccgcggggattataaac	REV352	Aatcagaatagctgatgtagtctcagtg
P353	FOR353	cgcggggattataaaccg	REV353	Tgtaatcagaatagctgatgtagtctca
R354	FOR354	gggattataaaccggggcg	REV354	Gggtgtaatcagaatagctgatgtagtc
G355	FOR355	gattataaaccgggcgcg	REV355	gcggggtgtaatcagaatagctg
D356	FOR356	tataaaccgggcgcgg	REV356	cccgcggggtgtaatc
Y357	FOR357	aaaccgggcgcgg	REV357	atccccgcggggtg
K358	FOR358	ccgggcgcggtc	REV358	ataatccccgcgggg
P359	FOR359	ggcgcggtcggtaaagt	REV359	tttataatccccgcgggg
G360	FOR360	gcggtcggtaaagttgttcc	REV360	cggtttataatccccgcg
A361	FOR361	gtcggtaaagttgttccatttttg	REV361	gcccggtttataatccccg
V362	FOR362	ggtaaagttgttccatttttgaagc	REV362	cgcgcccggtttataatc
G363	FOR363	aaagttgttccatttttgaagcg	REV363	gaccgcgcccgg
K364	FOR364	gttgttccatttttgaagcgaag	REV364	accgaccgcgccc
V365	FOR365	gttccatttttgaagcgaagg	REV365	tttaccgaccgcgcc
V366	FOR366	ccatttttgaagcgaaggttg	REV366	aactttaccgaccgcgc
P367	FOR367	ttttttgaagcgaaggttgtg	REV367	aacaactttaccgaccgcg
F368	FOR368	tttgaagcgaaggttgtggatc	REV368	tggaacaactttaccgaccg
D520	FOR520	gaagtaccgaaaggtcttaccgg	REV520	cacaaacacaactcctccgc
E521	FOR521	gtaccgaaaggtcttaccggaaaac	REV521	gtccacaaacacaactcctccg
V522	FOR522	ccgaaaggtcttaccggaaaac	REV522	ttcgtccacaaacacaactcc
P523	FOR523	aaaggtcttaccggaaaactcg	REV523	tacttcgtccacaaacacaactcc
K524	FOR524	ggtcttaccggaaaactcgac	REV524	cggtacttcgtccacaaacac
G525	FOR525	cttaccggaaaactcgacgc	REV525	tttcggtacttcgtccacaaac
L526	FOR526	accggaaaactcgacgc	REV526	acctttcggtacttcgtccac
K543	FOR543	aagggcggaaagtccaaattg	REV543	ggcctttatgaggatctctctgatt
K544	FOR544	ggcggaaagtccaaattgtaaaat	REV544	cttggcctttatgaggatctctc
G545	FOR545	ggaaagtccaaattgtaa aatgtaactg	REV545	cttcttggcctttatgaggatctc
G546	FOR546	aagtccaaattgtaaaatgtaactggatc	REV546	gcccttcttggcctttatgag
K547	FOR547	tccaaattgtaaaatgtaactggatcc	REV547	tccgcccttcttggc
S548	FOR548	aaattgtaa aatgtaactggatccgg	REV548	ctttccgcccttcttggc
K549	FOR549	ttgtaaaatgtaactggatccggc	REV549	ggactttccgcccttcttg
L550	FOR550	taaaatgtaactggatccggctg	REV550	tttggactttccgcccttc

Sequence of primers used for mutagenesis. Primers are listed including the name of the primer and its corresponding sequence.

2.2. General Molecular Biology and Recombinant DNA Methods

2.2.1. DNA Sequencing

All colonies of interest were sequenced at the Cardiff University DNA sequencing core. Sample preparation and submission was carried out as to the specifications found online at http://probe.biosi.cf.ac.uk/seq/submitting_samples.php with primers specified previously (see 2.1.5. *Sequencing*). Chromatograms returned were analysed utilizing MacVector/ Serial Cloner.

2.2.2. Purification of Plasmid DNA

2.2.2.1. From *E. coli* cell cultures

Plasmids were purified from <5ml of BL21 *E.coli* bacterial culture utilising kits obtained from Promega (Madison, WI, USA) Ltd as according to the manufactures protocol. The kit employs alkaline lysis of pelleted cells followed by 1-3M solution of sodium acetate, pH 5. DNA was eluted from the column using warm Milli-Q water. DNA concentrations and degree of purity were estimated by applying 1µl volumes to the NanoDrop® ND-1000 UV-Vis spectrophotometer (Thermo Fisher Scientific, MA, USA). When required, plasmid DNA was concentrated via use of a Speed Vac as per manufacturer's protocol.

2.2.2.2. From agarose gel

Purification of DNA from agarose gel slices was performed using the NucleoSpin Gel and PCR Clean-up kit (Machery-Nagel). The DNA was subsequently washed and eluted from the column using warm Milli-Q water. When required, plasmid DNA was concentrated via use of a Speed Vac as per manufacture's protocol.

2.2.3. Agarose Gel Electrophoresis

DNA fragments were separated and analysed based on size via agarose gel electrophoresis. Agarose (1% (w/v)) was suspended in TAE buffer (40mM Tris-acetate pH 9.5 and 1mM EDTA) and melted before the addition with ethidium bromide (0.5µg/ml). A 1/6 dilution of a 6x stock solution loading buffer (Tris pH 8.0, 40% sucrose (w/v), 0.01% (w/v) bromophenol blue) was added to DNA samples prior to loading onto the gel. Electrophoresis was then performed at between 80-120V for 30 - 60 mins in complementary tanks containing 1x TAE. DNA bands were visualised using an UV-transilluminator (GelDoc-It Imaging System, Ultra-Violet Products Ltd). For samples exceeding 20µl volumes, well casters were taped together to the desired volume.

2.2.4. PCR with GoTaq polymerase

GoTaq DNA polymerase (1.25 U) (Promega Ltd) was used to amplify DNA and was carried out in 50mM KCl, 10 mMTris-HCL pH 9.0, 0.1% (v/v) Triton X-100 and 1.5mM MgCl2 with 0.2mM dNTPs, 2M of each primer, template DNA (< 1-10 ng) made to a total volume of 50µl. Reactions were placed in a TC-412 thermocycler (Techne, MN, USA) and raised to 95ºC for 10 min. The following cycle was repeated 30 times: the temperature was taken to 95ºC for 30 sec for denaturation of the DNA: The temperature was dropped to 55-65ºC for 30 sec to allow the primers to anneal to the template DNA: The temperature was then taken to 72ºC for 1 min/kb of template DNA to allow DNA chain to elongate by the polymerase. After 25 cycles the temperature was held at 72ºC for 5 min to complete elongation before being held at 4ºC until removed from the thermocycler. The reaction (2µl) was analysed by agarose gel electrophoresis (1.0% (w/v)).

2.2.5. Site directed mutagenesis Phusion polymerase

Phusion High Fidelity DNA polymerase (1 U) (Thermo Fisher Scientific, MA, USA) was used to amplify DNA according to a modified Quikchange (Agilent Technologies, CA, USA) protocol in order to introduce trinucleotide deletions into the x11Fluc template and to

amplify the whole plasmid (Figure 3.8). Non overlapping complementary mutagenic primers (Table 2.2.) were designed to miss the amplification of a trinucleotide, corresponding to the amino acid targeted for deletion. Plasmids were amplified with mutagenic primers in a buffer containing 10mM Tris-HCL, pH 8.8, 50mM MgCl2 and 0.1% (v/v) Triton X-100 with 0.2mM dNTPs, 2µM of each primer, template DNA (<1-10 ng) made to a total volume of 50µl. Reactions were amplified in a TC-412 thermocycler as manufactures guidelines and raised to 98ºC for 30 sec. The following cycle was repeated 30 times: the temperature was taken to 98ºC for 15 sec for denaturation of the DNA: The temperature was dropped to 55-65ºC for 15 sec to allow the primers to anneal to the template DNA: The temperature was then taken to 72ºC for 30 sec/kb of template DNA to allow DNA chain elongation by the polymerase. After 30 cycles the temperature was held at 72ºC for 5 min to complete elongation before being held at 4ºC until removed from the thermo cycler. The reaction (2µl) was analysed by agarose gel electrophoresis (1.0% (w/v)). As per Quikchange protocol, digestion of all reactions was underwent via the addition of 1µl *Dpn1* directly to products of the SDM reaction and incubated for 1 hour at 37 ºC in order to remove adenomethylates *E.coli* (Dam+)-derived plasmids, that being template plasmids.

2.2.6. Restriction digestion

Dpn1 (1 U/µg DNA) (NEB) restriction digests were performed in 1 x CutSmart buffer (50mM potassium acetate , 20mM tris-acetate, 10mM magnesium acetate, 100µg/ml BSA, pH 7.9) in a total reaction volume of 50µl , and incubated at 37ºC (1 hr/µg DNA). *NdeI* and *NcoI* were heat inactivated at 65 ºC for 20 mins.

2.2.7. Ligation

DNA ligations were carried out using T4 DNA ligase (1µl/reaction) (NEB, MA, USA) in 1 x quick ligation reaction buffer (66mM Tris-HCL, 10mM MgCl2, 1mM DTT, 1mM ATP, 15% (w/v) PEG 6000, pH 7.6) (NEB,). Ligation of an insert gene into a plasmid was performed using 50 ng of vector DNA with a 3-fold molar excess of insert in a total

reaction volume of 20µl. Recircularisation (intramolecular) ligation of linear DNA was performed with 50ng of the linear plasmid. Ligation reactions were incubated at 4ºC overnight then the DNA was immediately transformed into competent cells.

2.2.8. Preparation of electro-competent cells

An LB broth (10ml) overnight culture was prepared from a single *E. coli* colony and incubated in a shaking incubator (200 rpm) at 37ºC overnight. The 10 ml overnight culture was diluted into two 500ml cultures of LB broth and grown to an OD_{600} of 0.4-0.8. The cells were harvested by centrifugation (1500 x g for 15 mins at 4ºC). The pellet was resuspended in 1L of ice-cold sterile water and harvested by centrifugation (1500 x g for 20 mins at 4ºC). The pellet was resuspended in 500ml of ice cold sterile water and the previous harvesting step was repeated. The pellet was resuspended in 250ml of ice-cold sterile water and harvested according to the last step. The pellet was resuspended in 100ml ice-cold sterile 10% glycerol and harvested as before. The pellet size was estimated and resuspended in an equal volume of 10% glycerol. The cells were divided into 40µl aliquots, snap frozen in liquid nitrogen and stored at -80ºC.

2.2.9. Transformation by electroporation

Electrocompetent cells stored at -80 ºC were thawed on ice. DNA (10ng – 20ng) was added to an aliquot of thawed cells (40µl), mixed and transferred to a pre-chilled, sterile electroporation cuvette and subjected to a 4.5–5 ms electrical pulse at 12.5 kV.cm-1 field strength using a gene pulser (Bio-Rad laboratories Ltd, CA, USA) with capacitance and resistance set to 25µF and 200 Ω, respectively. The cells were recovered by the addition of 460µl RT SOC in a sterile tube and incubated (37ºC at 200 rpm) for 1 hr. After recovery the cells were plated on LB agar supplemented with suitable antibiotics.

2.2.10. Transformation by heat shock

Chemical competent cells stored at -80^{o}C were thawed on ice. DNA (50ng -100ng) was added to an aliquot of thawed cells (40µl), mixed and incubated on ice for 20 minutes prior to inducing heat shock. The incubated cells were heated for 30 seconds in a 42°C water bath and immediately chilled on ice for an additional 2 minutes. The cells were recovered in 460µl SOC and incubated in a sterile tube and incubated (37^{o}C at 200 rpm) for 1 hr. After recovery the cells were plated on LB agar supplemented with suitable antibiotics.

2.2.11. Quantification of DNA

The NanoDrop spectrophotometer was routinely used to assess the DNA concentration of the FLuc variants.

2.2.12. Determination of concentration

Determination of the substrate concentrations were quantified routinely using UV/Vis spectroscopy and utilization of the Beer-Lambert law (molar extinction coefficient of 18 200 M^{-1} cm^{-1} for D-LH_2, from Morton et al. 1969) to derive the final concentration using the equation $A = \varepsilon Cl$.

2.2.13. Growth and Maintenance of *E.coli* strains

When required, bacteria from glycerol stocks were spread on fresh LB agar plates with appropriate antibiotics and grown overnight at 37^{o}C and single colonies picked from these and cultured in LB broth and antibiotic overnight at 37 ^{o}C, shaking circa 200rpm for further analysis. Where crude protein extracts were required 5ml cultures were grown overnight and centrifuged at 4000xg for 15 minutes. Plasmids were maintained short and long term in *E.coli* BL21. Aseptic technique was maintained throughout. Strains were stored at -80^{o}C as 0.5ml LB broth and antibiotic cultures with a final concentration of 8% (v/v) glycerol added to the *Ecoli* BL21 cells.

2.3. Methods for the Construction and Screening Single Amino Acid Deletion Variants

2.3.1. Cloning of 10x Histag x2FLuc DNA from pDEST17 into the pET16b backbone

A plasmid was constructed for the expression of x2 under the control of the pET16b in *E.coli* BL21 (DE3) pLysS to produce the vector pET16b-x2. The x2 luc gene was amplified from pDEST17 utilising primers to incorporate the restriction sites *Nde1* and *BamH1*. Both the linear fragment of gene encoding x2FLuc and pET16b-x11Fluc were digested with *Nde1* and *BamH1* restriction enzymes and following gel purification to isolate the vector was phosphorylated. The vector and insert were ligated using T4 DNA ligase and 5µl of the ligation reaction was transformed into NEB 5-alpha chemically competent cells and spread onto LB agar plates containing 100ug/ml amplicillin and incubated overnight at 37ºC. Sequencing was conducted on resultant colonies to confirm the presence of the x2FLuc gene within pET16b.

2.3.2. Methods for Screening

Colonies resulting from the primer-based site-directed deletion mutagenesis were screened using the primary colony screen and secondary colony screening procedures described. Colonies were optimally plated to provide circa. 100-500 colonies per petri dish after incubation overnight at 37ºC. This ensured colonies were not confluent as identification of mutants was easier.

2.3.2.1. Primary screening

The primary screen of colonies refers to the initial transformation of mutants resulting from the deletional mutagenesis strategy in which photographs are taken following bioluminescence induction. x11FLuc deletion mutants within *E.coli* BL21 (DE3) transformants where expression was induced by adsorbing colonies onto Hybond^{TM}- N nitrocellulose membrane (Amersham Biosciences Corp). Sterile tweezers were utilized to remove the colonies from the old medium onto fresh LB agar plates containing 100ug/ml

ampicillin and 1mM IPTG, on which they would be incubated between 3-6 hours at RT. The membrane was then transferred from these plates and placed face up onto a petri dish and bioluminescence was induced by spraying colonies with 500mM D-LH_2 in citrate buffer, pH5. Bioluminescence emitted by the colonies was collected using a charge-coupled device (CCD) camera from Nikon with 30s integration time in a dark room.

2.3.2.2. Secondary screen

After primary screening, colonies of interest were subjected to a secondary screen. The secondary screen refers to testing of the mutants with regards to apparent resistance to thermal inactivation. Single colonies of x11FLuc deletion mutants within *E.coli* BL21 (DE3) transformants were picked from plates using sterile pipette tips and spotted into rows of 4-5 genetically identical colonies onto LB agar along with x11FLuc template and incubated overnight at 37^{o}C. The induction of bioluminescence was similar as described in section 2.3.2.1 however colonies were incubated for 30minutes at 42^{o}C following induction with 1mM IPTG but prior to screening with LH_2. From the secondary screen, potential mutants were selected, cultured, and DNA prepared for sequencing to confirm trinculeotide deletions.

2.3.2.3. Bioluminescent Spectra from Whole Cells

To determine differences in the bioluminescent spectra from the variants generated, single colonies were picked using pipette tips and use to inoculate 5ml LB broth containing the appropriate antibiotic. Cells were grown at 37^{o}C overnight and subsequently diluted to an OD_{600} of 0.8-0.9 at which point they were induced with IPTG (1mM) for 3 hours. Following on from this, the cell suspension was mixed with a solution of 0.1M sodium citrate buffer pH5.0, containing 500mM D-LH_2 and spectra recorded within the Varian Cary Eclipse (Agilent Technologies, CA, USA).

2.3.2.3. Preparation of the 96-well format screening

Single colonies of deletion mutants were isolated and used to inoculate 5ml LB broth containing the appropriate antibiotic and grown overnight at 37^{o}C. Cultures were subsequently diluted to an $O.D_{600}$ of 0.8- 0.9 for induction with IPTG (1mM) for 3 hours. Following 3 hours induction, the colonies were diluted to the same absorbance at OD_{600nm}

and aliquoted 100µl into a 96 well plate in triplicate. The cells were harvested in the plates after centrifugation at 4500RPM for 20 minutes and any remaining LB broth was removed. The cell pellets were further resuspended in a mixture of 100mM TEM, 0.1% Triton-X100 and 10%glycerol. These were then stored at –80^{o}C.

2.3.2. Assaying with the 96 well format screening

Once ready for assaying, plates stored at -80^{o}C were removed from the freezer and left to come to RT. Each crude cell lysate was then assayed against several conditions. These conditions included; saturating conditions of the substrates at RT and following incubation at 42^{o}C and 60^{o}C, saturating conditions of the substrates pending a 5 minute settle time, saturating conditions of the substrates in the presence of 2mM PPi and non-saturating conditions of the LH_2.

2.4. Methods for Overexpression and Purification of Proteins

2.4.1. Overexpression of luciferases and mutants

Single colonies of each FLuc variant were picked and used to inoculate a LB broth (20ml) medium supplemented with 50µg/ml ampicillin. The culture was left to grow for circa 8 hours at 37°C, shaking (200rpm) and pending this, 0.5ml of the culture was placed into a further 100ml LB broth, under the previous conditions, overnight. Subsequently, 20ml of the culture was used to inoculate an additional 400ml LB broth with antibiotics and grown at 37°C for circa 1-2 hours. The growth of the culture was monitored until the cells measured an OD_{600} between 0.6-0.7 AU by a spectrophotometer (Pharmacia Biotech, Sweden) at which point IPTG, at a final concentration of 1mM was used to induce protein expression. The cultures were reincubated at RT, shaking (200rpm) for a further 6 hours. Pending this, the cells were harvested at 14700xg for 30 minutes at 4°C and the supernatant discarded.

2.4.2. Cell Lysis and Purification of Variants

Protein purification was carried out on a HisPur Ni-NTA affinity column (ThermoFisher Scientific, MA, USA). Bacterial pellets of overexpressed variants were resuspended in 5ml lysis buffer (Buffer A supplemented with 2% Triton X-100 (v/v), 20mM imidazole) per gram of pellet and incubated with 10ul benzonase nuclease (250U/ul) on ice for 15-30minutes. Buffer A comprises 10mM phosphate, 2.7mM KCl, 0.3M NaCl, 10mM 2β-mercaptoethanol, 20% glycerol (v/v), 1 xEDTA-free protease cocktail inhibitor (Roche Diagnostics, USA). The suspension was centrifuged at 12000rpm for 30 minutes at 4°C and the supernatant removed as the soluble fraction. 10ml volumes of supernatant comprising the soluble fraction was loaded onto a HisPur Ni-NTA (pre-equilibrated in Buffer A supplemented with 20mM imidazole (IMD) (ThermoFisher Scientific, MA, USA) and the flow through was reapplied 1-2 times prior to 50mM imidazole in Buffer A application to the column in order to elute non-specifically bound proteins. His tagged mutants were then eluted from the column in 2.5ml sets of 500mM IMD in Buffer A. 20µl of elution were

kept on ice and assayed for bioluminescence. Immediately fractions containing activity were desalted by applying them to PD-10 columns, pre equilibrated in 25ml storage buffer.

2.4.3. Luminometric quantification during protein purification

The average activity of 1µl of each fraction assayed by addition of 500mM D-LH_2 and 1mM ATP in TEM (pH7.8) in triplicate. The concentrations utilized are approximately 10x the calculated K_M values for x11FLuc.

2.4.4. Quantification of Protein Concentration

Bradford assays (Bradford, 1976) were carried out according to the manufacturer's protocol. 5µl of protein sample or bovine serum albumin (BSA) standards were applied in triplicate to a 96 well plate and 200µl of protein assay reagent was added. The reaction was left to incubate at RT for 10 mins prior to reading. Absorbance was measured at 595nm in triplicate and protein concentration determined by linear regression of the standard plot.

2.4.5. SDS-PAGE of Expression of Variants

To analyse the protein expression from the FLuc variants, samples (10µl) were prepared and loaded onto a sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) comprising a 12.5% separating gel and stacking gel. The gels were submerged in an SDS running buffer (Glycine 144g/2L, Tris Base 40g/2L, 0.1% (w/v) SDS). Once the broad range pre-stained protein marker (NEB) and samples were loaded, 200V was applied to the gel for 40-60 minutes until separation of the observed molecular bands was sufficient.

2.4.6. Staining of the SDS Gel

Once the SDS PAGE had run to completion, the stacking gel was discarded and the separating gel was placed into a container a Coomassie staining solution (40% Methanol (v/v), 10% Acetic acid (v/v), 0.1% Coomassie (w/v)). The gel was then incubated for 24

hours on the orbital shaker (New Brunswick Scientific, USA). After staining the Coomassie stain was removed and replaced with a destaining solution (40% Methanol (v/v), 10% Acetic acid (v/v)) for 48 hours.

2.5. Firefly Luciferase Characterization Methodologies

2.5.1. Luminometric Methods

All bioluminescent measurements were carried out in a Fluoroskan Ascent (Thermo Scientific, MA, USA) instrument utilizing one injector in order to inject the substrates of the reaction into wells within a 96-well plate containing the FLuc at the required concentration, diluted to 10µM in pre-chilled TEM buffer. Measurements were integrated for 20ms-1s over variable time periods comprising between 50 to 250 consecutive measurements. Approximate K_M values were deduced for each enzyme and concentrations were utilized representing 10-times above and below the calculated K_M. Additionally, a correction was made in order to overcome the sensitivity of the photomultiplier tube (PMT) to different wavelengths of emission by measuring the fluorescence spectrum of lucifer yellow-CH (Sigma-Aldrich, MO, USA) and correcting it to an absolute spectrum supplied by the manufacturer.

2.5.2. Measurement of Bioluminescent Spectra

Bioluminescent spectra were calculated with a Cary Eclipse Fluorescence Spectrophotometer. The fluorimeter and associated software was set up with variable emission slit whilst the PMT setting was set to 950V. All data was corrected against the calibration made for the PMTs sensitivity towards different wavelengths. In addition, all data was corrected against the baseline. In all experiments, saturating conditions were set up that substrates D-LH_2 and ATP were 10x the K_M for their substrates respectively.

2.5.3. Determination of Kinetic Constants

Since the bioluminescence reaction displays flash kinetics, conventional methodology may not be employed in order to determine the kinetic constants. However, the peak intensity (I_{max}) is a representation of the pre-steady state of maximal light intensity at a given

substrate concentration and occurs at one turnover of the enzyme (Brovko *et al*, 1994). Use of I_{max} to determine kinetic constants obeys Michaelis-Menten kinetics for the reaction prior to other complicating factors such as end product inhibitors (Ugarova, 1989). The Imax of the different FLucs were measured as described previously in a Fluoroskan Ascent instrument. Measurements to determine kinetic constants with regards to LH_2 were made in conditions where the concentration of ATP was saturating whilst the concentration of LH_2 were varied. Measurements to determine kinetic constants with regards to ATP were made in conditions where the concentration of LH_2 were saturating whilst the concentration of ATP were varied. 50µl of solutions containing the varying concentrations of the substrate (LH_2, ATP) diluted in TEM (pH7.8, RT) were injected onto 0.5µM Luc. Concentrations of D-LH_2 were diluted such that the final working concentration within the reaction mix included 0.1, 0.5, 1, 5, 10, 20, 35, 70, 140, 200µM. Concentrations of ATP were diluted such that the final working concentration within the reaction mix included 0.1, 0.5, 10, 25, 50, 100, 200, 400, 800, 1000µM. Once injection had occurred, measurements were taken immediately recorded over 1 minute. All measurements were made in triplicate for each concentration point and each experiment was repeated independently between 2-4 times for every FLuc with all substrates. Data were analysed by plotting substrate concentration against Imax and subsequently kinetic constants were derived using the Hanes-Woolf plot (Athel-Cornish Bowden, 1999).

2.5.4. Specific Activity Determination

Measurements were made to determine specific activities of the enzymes in saturating conditions of both substrates (500µM LH_2, 1mM ATP). 50µl of solutions containing saturating concentrations of the substrates were injected onto 50µl of 0.5µM Luc at which point measurements were integrated over 1s for a total period of 250s minutes. All measurements were made in triplicate for each FLuc.

2.5.5. pH Dependence of Bioluminescent Spectra

Measurements were made to determine spectral changes of the enzymes within buffers at variable pH (pH6.3, 6.8, 7.3, 7.8, 8.3, 8.8) whilst saturating conditions of both substrates (500µM Luc, 1mM ATP). TEM was utilized as buffer and pH was adjusted using acetic acid or sodium hydroxide. 50µl of solutions containing saturating concentrations of the substrates at variable pH were injected onto 50µl of 0.5µM Luc and measurements were made within a fluorimeter. All measurements were made in triplicate for each FLuc.

2.5.6. Determination of pH Dependence of Activities

Measurements were made to determine activities of the enzymes within buffers at variable pH (6.3, 6.8, 7.3, 7.8, 8.3, 8.8) whilst saturating conditions of both substrates (500µM D-LH_2, 1mM ATP). 50µl of solutions containing saturating concentrations of the substrates at variable pH were injected onto 50µl of 0.5µM Luc and measurements were made within a fluorimeter. All measurements were made in triplicate for each FLuc.

2.4.7. Determination of Thermal Stability

Typical working temperatures for applications involving luciferase range from between room temperature to 50^{o}C as such, characterization of mutants up to this temperature is desirable. Determination of thermal stability was carried out using two methods. Firstly, 500µl aliquots of 0.5µM Luc solutions at pH7.8 at 4^{o}C were incubated at either 20^{o}C, 30^{o}C, 35^{o}C, 40^{o}C, 45^{o}C, 50^{o}C and 150µl aliquots removed every 15 minutes over a total period of 1 hour. All removed aliquots were placed onto ice for 30 minutes prior to assay conducted at room temperature. The second method investigated the activity of the mutants at a set temperature. For this assay, 500µl aliquots of 0.5µM Luc solutions at pH7.8 were pre-incubated at either 30^{o}C, 35^{o}C, 42^{o}C for 10 minutes prior to assaying at a temperature equal to incubation. For both methods, measurements were made within saturating conditions of both substrates (500µM LH_2, 1mM ATP) at pH7.8 injected onto 0.5µM Luc. Measurements

were made within a luminometer with point measurements integrated over 0.2s for a total period of 5s. All measurements were made in triplicate for each FLuc

Chapter 3

Construction and Screening of Single Amino Acid Deletion Mutants Within Thermostable and pH tolerant *Photinus Pyralis* x11 Firefly Luciferase

3.1. Chapter Summary

In this work I hypothesise that the incorporation of single amino acid deletions within FLuc derived from *Ppy* will exhibit potentially beneficial characteristics upon the protein as has been shown in previous studies with other model proteins such as; the green fluorescent protein (GFP) and TEM 1 β lactamase. Conventionally, terminals and loops are more tolerant to single amino acid deletions in comparison to more highly structured elements such as alpha helices and beta sheets. Therefore, molecular visualisation was utilised to identify loop regions within the structure of Fluc which would be interesting targets for deletion. Utilising a primer based targeted strategy, single amino acid deletion mutants of Fluc were designed and the resulting effects upon bioluminescence activity, colour of emission and apparent stability were assessed. x11 luciferase is remarkably tolerant to single amino acid deletions within its structure in respect to bioluminescence activity and in terms of ability to retain the of characteristics of native x11 luciferase, such as resistance to thermal inactivation and bioluminescence emission peak wavelength. Interestingly, through this deletional strategy a variety of colour mutants were obtained, displaying significant spectral shifts with large variations in the degree of red-shift. In addition to this, it appears that certain single amino acid deletions improve bioluminescence activity compared to x11 luciferase, which will be explored in further chapters.

3.2. Introduction

An overriding dogma in protein engineering is that amino acid deletions tend to compromise the structural integrity of proteins, abolishing function however more recent studies have shown that this is not always the case (Simm *et al.*, 2007; Baldisseri *et al*., 1991; Arpino *et al.*, 2014). Simm *et al*. (2007) used deletion mutations to enhance the activity of TEM-1 ß-lactamase towards a ceftazidime substrate, with a 64-fold increase in activity of a deletion in P174 compared to the wildtype (WT). Similarly, deletions of six amino acids promoting a closer compacted state within *Staphylococcal* nuclease increases the catalytic activity and stability compared to the WT (Baldisseri *et al*., 1991). In addition to this, the insertion/deletion mutations of single amino acids commonly occur throughout nature (Taylor *et al.*, 2004) with small insertions most often tolerated in proteins (Shortle and Sondek, 1995).

FLuc has been heavily engineered in terms of a number of desirable characteristics, e.g. activities, kinetics, emission colours and resistances to thermal inactivation. To date, this has been achieved using substitution mutations, insertion of amino acids and domain replacement. However, no attempts have been made to incorporate and test single amino acid deletions in FLuc, which would increase the sample space and improve the chances of isolation of novel phenotypes. FLuc provides a good model protein in which to explore this rationale, and that is the aim of this thesis.

A powerful screening strategy is critical in any FLuc protein engineering effort, the most common strategy employed was that as first described by Wood and Deluca (1987). This approach required *E.coli* colonies expressing functional FLuc to be saturated with a citrate buffer containing D-LH_2 at pH 5.0 since it was discovered that whilst LH_2 cannot be efficiently delivered into prokaryotic cells at physiological pH under acidic conditions, such as that managed by a citrate buffer at pH 5, LH_2 readily passes through the cell membrane (Wood and Deluca, 1987; Jawhara and Morodn 2004). This change can be attributed to the carboxyl group of the molecule being charged at physiological pH whereas at the lower pH, the LH_2 is protonated which allows for the passage into cells. Since this time, others have made adaptations to

improve upon this initial approach (Erica Law, PhD Thesis, University of Cambridge, 2004) by incorporating:

- The introduction of a secondary screen in addition to the primary screen (single colonies).
- The additional screening of mutants after incubation at 42^{o}C.
- The analysis of the light emission from colonies by image analysis software.
- Subsequent screening of cell lysates in a 96-well or 384-well formats.

The aim of this investigation is to incorporate sequential single amino acid deletions via a primer design based method within both conserved and non-conserved regions of thermostable *Ppy* x11FLuc, which has stable light emission characteristics up to 50^{o}C, to identify whether its phenotype is tolerant to the deletions in internal loops. This chapter also seeks to identify whether deletions can not only retain but further improve on the already beneficial characteristics displayed by x11FLuc, such as activity or resistance to thermal activation in order to design an enhanced protein developed by both substitution and deletion methodologies. I also aim to determine whether deletions appear to promote any desirable or potentially desirable novel phenotypes.

3.3. Results and Discussion

Initially attempts were made to utilise the Tri-NEX transposon based strategy in order to generate a library of luciferase single amino acid deletion mutants with the intention of utilising a directed evolution approach throughout screening however this proved unsuccessful and as such a primer design method was employed. The advantage of using a targeted approach to incorporate single amino acid deletions is the ability to develop a rationale to target interesting regions within the protein and incorporate specific mutations within these regions.

3.3.1. Analysis of Secondary Structure

Secondary structure elements such as the loops within a protein are more tolerant to deletion than more ordered structures such as □-helices and □□sheets (Arpino *et al*., 2014). In order to identify these regions more likely to exhibit beneficial characteristics as a result of deletions, a secondary structure of the x11FLuc was required. To date, no structural determination of x11FLuc has been conducted to access the degree of structural variation imparted as a result of the 11 mutations that have been incorporated into the WTFLuc. YASPIN secondary structure prediction software (Centre for Integrative Bioinformatics, University of Amsterdam) was assessed for utility by comparison of submission results obtained for the WTFLuc to those results obtained from *Ppy* FLuc protein structure from Protein Data Bank (PDB) file *Ppy* FLuc ID: 4G36 as resolved to 2.62Å (Sundlov *et al.,* 2012) (Table 3.1.).

Results derived from this comparison highlighted that the YASPIN secondary structure prediction identified secondary structures correlated to the secondary structure displayed in the PDB model, however there were some discrepancies between the prediction and the model such as between positions 300 – 315, the model presented α-helix and a β-sheet whilst the prediction presented an α-helix between 306 – 315, unable to identify the β-sheet. In addition to this, the prediction was able to identify the secondary structure, however, it was unable to determine the exact boundaries of the structural element.

YASPIN secondary structure prediction was then conducted with the known x11FLuc sequence in order to determine whether it was likely that structural changes had occurred within x11FLuc as a result of the x11 mutations. YASPIN was only able to identify two potential sites that had been altered compared to the WTFLuc, those being, a deletion of a β-sheet at position 361-363 and the shortening of a β-sheet at position 390-395.

These results suggest that the mutations present within x11FLuc do not affect the structure of the protein significantly. In order to identify loop regions present within x11FLuc, the secondary structure from the crystal structure of *P. pyralis* luciferase reported by Sundlov *et al.,* (2012) Protein Data Bank (PDB) file *Ppy* FLuc ID: 4G36 as resolved to 2.62Å (Figure 3.1.) was utilised.

DSSP (Kabsch and Sander, 1983) and STRIDE (Frishman and Argos, 1995) analysis conducted on the *Ppy* FLuc ID: 4G36 structure identified that approximately 29% and 21% of the total protein was comprised of helical and □-sheet elements respectively, suggesting that remaining elements (including turns and loops) comprise circa 50% of the total protein. However, direct observation of the crystal structure showed that there were a total of 40 loops ranging in size from 1 amino acid in length to 15 amino acids within *Ppy* FLuc. The loops comprise a total of 224 amino acid residues, equal to 40% of the residues.

All of these positions are potential candidates for deletion however within the remit of study it is not experimentally feasible to introduce deletions at all these positions. Instead, in this work, candidate positions based on homology (degree of conservation), secondary structure and 3D-structural analyses (i.e. those that are likely to impact on important active site residues) were targeted for deletion.

Table 3.1. Comparison of Secondary Structures both Known and Predictions made by YASPIN.

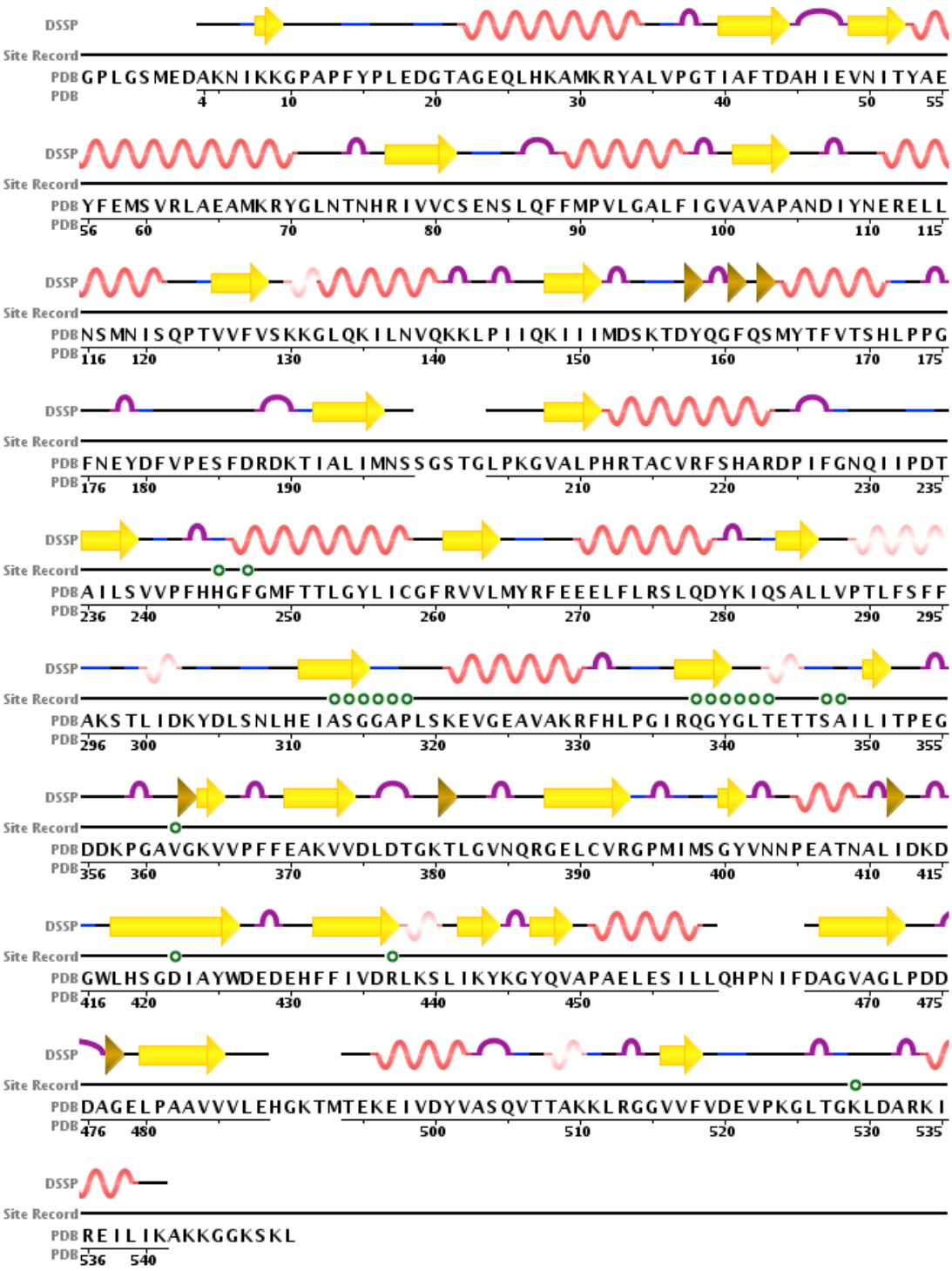
WTFLuc Known Secondary Structure	**WTFLuc YASPIN Prediction**	**x11FLuc YASPIN Prediction**
H 22 – 34	H 22-33	H 22-33
E 40 – 44	E 40-45	E 40-45
E 49 – 52	E 49-51	E 49-51
H 53 – 70	H 54-70	H 54-70
E 77 – 81	E 78-83	E 78-83
H 89 – 97	H 88 – 98	H 88 – 98
E 101 – 104	E 102 – 104	E 102 – 104
H 111 -121	H 112 – 120	H 112 – 120
E 125 -128	E 124 – 131	E 124 – 131
H 130 – 140	H 133 – 144	H 133 – 144
E 148 – 151	E 149 – 154	E 149 – 154
H 164 – 171	H 164 – 173	H 164 – 173
E 192 – 196	E 194 – 199	E 194 – 199
E 208 – 211	E 210 – 213	E 210 – 213
H 212 – 223	H 215 -230	H 215 -230
E 236 – 239	E 237 – 240	E 237 – 240
H 246 – 258	H 243 – 260	H 243 – 260

WTFLuc Known Secondary Structure	**WTFLuc YASPIN Prediction**	**x11FLuc YASPIN Prediction**
E 261 – 264	E 263 – 266	E 263 – 266
H 270 – 279	H 272 – 281	H 272 – 281
H 270 – 279	H 272 – 281	H 272 – 281
E 284 – 286	E 284 – 288	E 284 – 288
H 289 – 295	H292 – 299	H292 – 299
H 300 – 302	H 306 – 315	H 306 – 315
E 311 – 315		
H 321 – 330	H 323 -332	H 323 -332
E 337 -340	E 337 – 342	E 337 – 342
H 343 -354		
E 350 -351	E 350 -354	E 350 -354
E 364 – 365	E 361 – 363	
E 370 – 374	E 371 – 377	E371 – 37
E388 - 393	E389 – 395	E390 – 395
E 400 – 401	H 401- 405	H 401- 405
H 405 - 409	H 407 – 414	H 407 – 414
E 418 – 426	E 419 – 421	E 419 – 421

WTFLuc Known Secondary Structure	**WTFLuc YASPIN Prediction**	**x11FLuc YASPIN Prediction**
E 432 - 437	E 423 – 429	E 423 – 429
H 438 – 440	E433 – 438	E433 – 438
E 442 444	E 443 – 446	E 443 – 446
E 447 - 449		
H 451 458	H 453 – 461	H 453 – 461
E 467 - 472	E 466 – 473	E 466 – 473
E 480 - 485	E 482 – 489	E 482 – 489
H 495 -502		
H 508 -510		
E 516 -518	E 516 – 522	E 516 – 522
H535 - 539	H 531 - 545	H 531 - 545
E 516 -518	E 516 – 522	E 516 – 522
H535 - 539	H 531 - 545	H 531 - 545

Comparison of WTFLuc secondary structure as determined by YASPIN secondary structure prediction software against known structure of WTFLuc. Positions of α-helices and β-sheets present within luciferase as determined by YASPIN and the corresponding positons of α-helices and β-sheets within the known structure are compared. The x11FLuc YASPIN prediction is also included to compare differences in secondary structure as a result of the 11 mutations within x11FLuc. H: α helix, E: β sheet.

Figure 3.1. Secondary Structure of *Ppy* FLuc ID: 4G36



Schematic of Secondary Structure derived from *Ppy* FLuc ID: 4G36 highlighting the position and type of secondary structural elements. Analysis identifies that the loops comprise between 40 -50% of the protein whilst approximately 29% and 21% of the total protein was comprised of helical and □-sheet elements respectively. Figure sourced from PDB, Sequence Chain View, 4G36.

3.3.2. Regions of Disorder

From both secondary structure and alignment analysis a number of potential sites for deletion were identified, however, further analysis was required in order determine the specific sites of interest.

Proteins in their native state have regions that do not adopt a single conformation, resulting in flexibility. This flexibility can impart useful traits such as interaction with binding partners or to be involved in molecular recognition of biological processes such as regulation, signaling and cell cycle control (Dunker *et al.,* 2002; Wright and Dyson, 1999). Following binding, commonly proteins will transition to an ordered, rigid structure (Ishida & Kinoshita, 2007).

It is considered that the inherent flexibility of highly disordered segments imparts instability within a protein promoting difficulties both in purification of the protein and in subsequent crystallisation (Oldfield *et al*., 2005). Therefore, for the purposes of structural genomics or purposes that require large volumes of correctly folded protein it is desirable to remove these regions and impart greater stability, providing these segments do not hinder the functioning of the protein. There are multiple prediction tools that have been developed for the purposes of disorder prediction such as those supplied by Robetta or PSIPRED and PrDOS (Ishida and Kinoshita, 2007). These methods utilise the amino acid sequence of a protein in combination with the tertiary structure of template proteins to identify disorder. Initially, a support vector machine (SVM) algorithm maps specific residues after which a secondary analysis assumes conservation of intrinsic disorder occurring within protein families and utilises this in PSI-BLAST against an index of disorder. The overall prediction is presented as a measure of both set of results and due to the rigorous nature of this test, PrDOS was utilised for the purposes of analysing the x11FLuc amino acid sequence to identify regions of disorder.

Figure 3.2. displays the output generated following x11FLuc amino acid sequence submission to PrDOS. The tool indicates that the N- and C-terminals are highly disordered indicating a greater than 90% chance that the primary 4 amino acid residues M1-A4 and the final 15 amino acid residues I535 –L550 are disordered compared to

surrounding regions. It is common for N- and C-terminals to display intrinsic disorder since they tend to lack the stabilisation and coordination often imparted by surrounding amino acid residues. Conventionally, it has been seen that termini are more tolerant to deletions than other structures within a protein (Arpino *et al*., 2014).

Indeed throughout a single amino acid deletion survey conducted within eGFP, even secondary structural elements such as □□helices occurring within terminals were tolerant to deletion. We can consider that this tolerance may be due to the removal of these unstable residues. Previous studies have explored deletions within the N- and C-terminals of luciferase however to date, single amino acid deletion mutagenesis has not been utilised (Sala Newby and Campbell, 1994; Sung and Kang, 1998; Wang *et al*, 2002). Therefore, I shall consider both x11FLuc N- and C-termini as the first candidate targets for the single amino acid deletion strategy.

In addition to the termini, there are another 6 amino acid residues P173 to Y179 that indicate a probability of disorder of more than 90% within the loop comprising L172-T191. This prediction suggests that this loop is unlikely to adopt a single conformation and as such should have a high degree of flexibility. The requirement for such flexibility is often considered to be vital in allowing some proteins to interact with binding partners or to play a role in biological events such as molecular recognition, signalling and cell cycle control (Dunker *et al* 200; Wright and Dyson, 1999). At present, this loop has no reported associations with such functions and it is reasonable to explore this loop further. Molecular visualisation using Visual Molecular Dynamics (VMD, Beckman Institute for Advanced Science and Technology, University of Illinois, USA) using the model provided by the crystal structure of *P*. *pyralis* luciferase reported by Sundlov *et al*. (2012) (PDB file *Ppy* FLuc ID: 4G36) which highlighted that this loop is located on the solvent exposed surface of the C-terminal domain (Figure 3.7). Therefore, it is interesting to identify whether this disordered loop is tolerant to deletions and therefore shall be chosen as the second set of candidates, along with the first candidates (N- and C-terminal loops) (Table 3.4.).

Figure 3.2. Disorder prediction output based on the x11FLuc Amino Acid Sequence

```
MEDAKNIKKGPAPRYPLEDGTAGEQLHKAMKRYAQVPGTIAFTDAHIEVNITYAEYFEMSVRLAEAMKRYGLNTNHRIVVCSENSLQFFMPVLGALFIGVAVA
XXXX------------------------------------------------------------------------------------------------------

PVNDIYNERELLNSMNISQPTVVFVSKKGLQKILNVQKKLPIIQKIIIMDSKTDYQGFQSMYTFVTSHLPPGFNEYDFKPESFDRDKTIALIMNSSGSTGLPK
------------------------------------------------------------------------XXXXXX----------------------
GVALPHRCACVRFSHARDPIFGNQIKPGTAILSVVPFHHGFGMFTTLGYLICGFRVVLMYRFEEELFLRSLQDYKIQSALLVPTLFSFFAKSTLIDKYDLSNL
------------------------------------------------------------------------------------------------------

HEIASGGAPLSKEVGEAVAKRFHLPGIRQGYGLTETTSAILITPRGDYKPGAVGKVVPFFEAKVVDLDTGKTLGVNQRGELCVRGPMIMSGYVNNPEATNALI
------------------------------------------------------------------------------------------------------

DKDGWLHTGDIAYWDEDEHFFIVDRLKSLIKYKGYQVAPAELESILLQHPNIRDAGVAGLPDDDAGELPAAVVVLEHGKTMTEKEIVDYVASQVTTAKKLRGG
------------------------------------------------------------------------------------------------------

VVFVDEVPKGLTGKLDARKIREILIKAKKGGKSKL
-----------------------XXXX-XXXXXX
```

Disorder prediction output following submission of the amino acid sequence of x11FLuc to PrDOS. PrDOS calculates the probability of disorder and denotes those amino acid residues with a 90% or greater of disorder and annotates these residues with X. Those residues indicating such disorder include the N-terminal M1-A4, the C-terminal I540 – L550 with the exception of K544 and residues P174 – Y179.

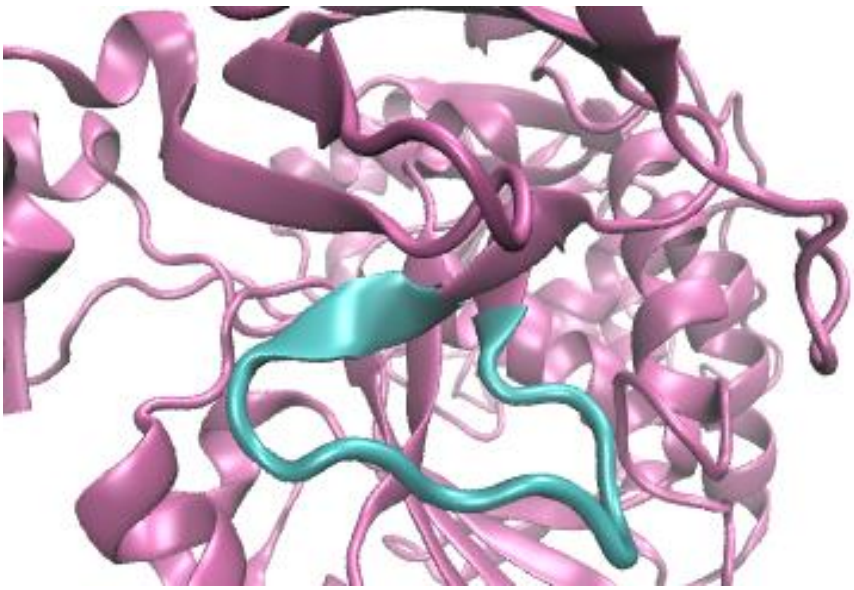
.

3.3.3. The Omega (Ω) Loop of Luciferase

Random mutagenesis experiments have identified a region within luciferase, substitution of which gives wide-ranging phenotypes dependent on the type of mutations that are targeted to this particular site. This "hot spot" is the loop T352-F368 in which mutations E354R and D356Y have been shown to confer resistance to thermal inactivation to x2FLuc and x11FLucs (White *et al.* 1996; Jathoul *et al.*, 2012). In addition to substitution mutations, Tafreshi *et al* (2007) noted that insertion of a critical residue Arg356 within this region increased the optimum temperature for bioluminescent activity and produced a bimodal bioluminescence spectrum red-shifting the bioluminescence emission peak to a wavelength maximum (λmax) of 615nm and a smaller peak at 560nm. This study additionally determined that whilst these changes occurred as a result of the insertion, the basic kinetic properties of the enzyme had been retained. Furthermore, Moradi *et al* (2009) confirmed this finding of red shifts as a result of an insertion within this region and sought to identify the role charge distribution played in determining the colour of bioluminescence. Both Arg356 and Lys356 conferred red bioluminescence whist Glu356, an amino acid with a negative side chain and Gln356, an amino acid with a neutral side chain presented no such shift.

Considering the apparent plastic nature of this loop, molecular analysis was conducted to explore the usefulness of single amino acid deletions within this loop, as an additional target for deletional mutagenesis. Molecular analysis was conducted within VMD using the model provided by the crystal structure of luciferase reported by Sundlov *et al*. (2012) (PDB file FLuc ID: 2D1Q). Figure 3.3. shows the loop as it exists within the native state of the protein and it is interesting that this loop conforms to what is considered in protein structural terms as an omega (Ω) loop motif, despite not as yet having been defined within luciferase. Ω loops are non-regular secondary structure elements characterised by a polypeptide chain that follows a loop-shaped course in three dimensional space, giving rise to, and as so named, an Ω loop (Fetrow *et al.*, 1995). Since 1986 when the structure was first described, it has become clear that these relatively simple structures are often far more dynamic and

Figure 3.3. Identification of the Omega Loop of Luciferase



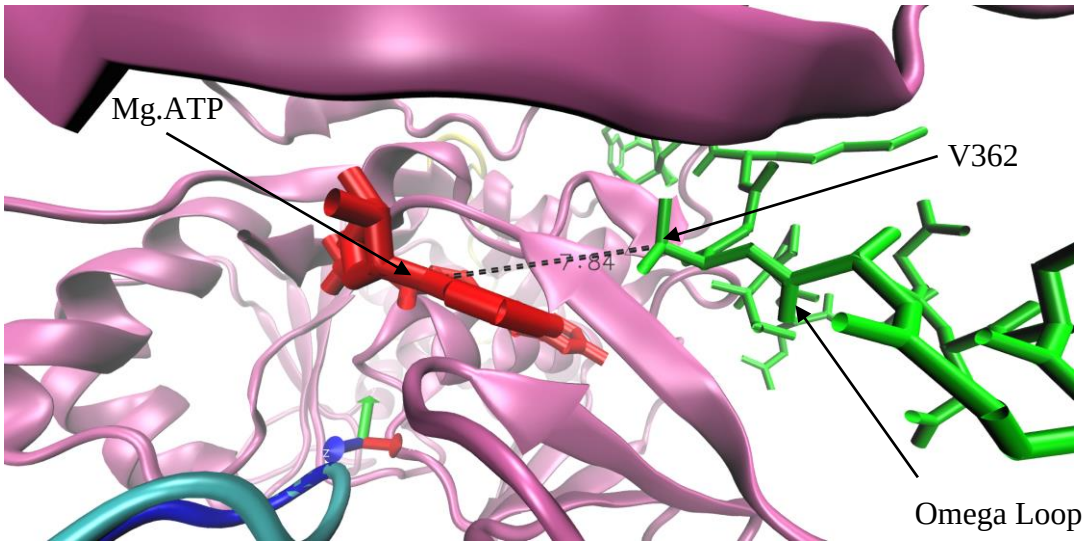
The characteristic shape of the omega loop within x11FLuc. Residues comprising the omega loop, T352-G363 highlighted in cyan. Image generated via VMD using PDB file ID: 4G36.

involved in both protein function, molecular recognition and are important for substrate specificity (Fetrow *et al.*, 1995). Characteristics of an omega loop include a large number of hydrogen bonds which are less periodic than seen in other structural elements and are found at the surface of the protein (Fetrow *et al.*, 1995). Taking this into account, further molecular visualisation of *Ppy* FLuc ID: 4G36 in which the luciferase is locked in to its adenylate-forming conformation, showed that the Ω loop of luciferase lies adjacent to the active site of the protein (within circa 10Å) and may play a role in the function of luciferase, as seen in other proteins (Figure 3.4). As such, it is interesting to identify whether this Ω loop is tolerant to deletions, since both substitutions and insertions within this region show large changes in activity. This will be the first example of the detailed exploration of the function of the Ω loop in the bioluminescent reaction. Therefore, the loop comprising T352-F368 was targeted for single amino acid deletion as a third candidate in this study (Figure 3.7) (Table 3.4).

3.3.4. Molecular Graphics Analysis to Identify Regions within 5 Å of the Active Site

From previous studies within other model proteins it has been identified that substitutions within the proximity of the active site can alter function (Kaiser *et al.*, 1988). It is common for protein engineering attempts to focus on amino acid residues within or proximal to the active site when screening for mutants with altered catalytic activities. Recently, single amino acid deletions close to the tripeptide chromophore of GFP have been shown to inhibit activity (Arpino *et al.*, 2014). In this work, a tolerance survey was implemented to identify which single amino acid deletions retained or lost fluorescence within GFP. This led to the conclusion that fluorescence was better retained or even enhanced, when the deletion occurred furthest from the chromophore, whereas deletions occurring proximal to the chromophore led to loss of fluorescence or reduced chromophore maturation rates. GFP is highly structured protein in which the chromophore microenvironment provided by the □□barrel is critical for activity preventing the quenching of the fluorescence by water molecules (Tsien, 1998). In such a highly organised protein it is not surprising that there is little flexibility in response to the registry shifts that can occur as a result of a single amino acid deletions (see Section 1.2.3.). In contrast, x11FLuc is roughly twice as large as GFP

Figure 3.4. Proximity of MgATP to Resdiue V362 within the Omega Loop



The omega loop within 10Å of the active site. The loop T352-F368, identified as an omega loop, is highlighted in green, within 7.84Å of MgATP in red. Image generated via VMD using PDB file ID: 2D1Q

and is a complex flexible, dual-domain globular enzyme. As previously discussed (see Section 3.3.1), according to analysis provided by DSSP and STRIDE, 40%-50% of FLuc comprises secondary structural elements in comparison to GFP of which 60% comprise secondary structural elements. This being the case, it is of interest to survey the tolerance exhibited by x11FLuc to single amino acid deletions occurring proximal to the active site to assess whether more globular proteins promote an increase in the level of tolerance to deletion within this site.

To assess if candidate loops were present within x11FLuc 5Å of the active site (Branchini *et al.,* 2003), an analysis within VMD was conducted using the model provided by the crystal structure of *P*. *pyralis* luciferase reported by Sundlov *et al*. (2012) (PDB file *Ppy* FLuc ID: 4G36) (Table 3.2.). The analysis identified a total of 38 amino acid residues within 5Å of the active site comprising loops, □-helices and □□sheets. Within the remit of this investigation the primary focus was to target loops in or near the active site to isolate novel phenotypes. Two loops of interest were identified which comprised more than 7 amino acids those being, T352-F368 (noted previously see Section 3.3.3.) and D520-L526, which are 12 and 14 amino acids in length, respectively. Therefore, the loop comprising D520-L526 was targeted for single amino acid deletion as a fourth candidate in this study (Figure 3.7) (Table 3.4).

3.3.5. Multiple Sequence Alignment of Beetle Luciferases

Since the aim of this thesis is to use the deletional strategy as an alternative and/ or complementary method for FLuc engineering, it was important to survey the effect of deletions on bioluminescence activity within both conserved and non-conserved regions of the enzyme. It is likely that deletions will be less tolerated within conserved regions compared to non-conserved regions, but to date, this assumption has not been tested. Therefore, protein sequence alignment of different beetle luciferases (shown in Figure 3.5-3.6) was used to identify these regions of conservation with particular reference to the regions previously selected in prior sections of this chapter. Utilising MacVector Inc (North Carolina, USA) a total of 20 homologues of beetle luciferases whose sequences were

Table 3.2. Amino Acid Residues occurring within 5Å of the active site

Amino Acids	**Structure**	**Loop**	**Length of Loop Comprising Amino Acid**
S198	Loop	N197 – G207	6
R218	Alpha Helix		
H245-F247	Alpha Helix		
T251	Alpha Helix		
L286	Beta Sheet		
A313-G315	Beta Sheet		
G316-P318	Loop	G316 – L319	4
R337-Y340	Beta Sheet		
G341	Loop	G341	1
L342-E344	Alpha Helix		
T346	Alpha Helix		
S347-A348	Loop	S347 – A348	4
I351	Beta Sheet		
V362	Loop	T352-F368	12
S420	Beta Sheet		
D422	Beta Sheet		
I434	Beta Sheet		
R437	Beta Sheet		
T519-K530	Loop	D520-L526	14

Table highlighting all amino acid residues occurring within 5Å of the active site of FLuc. The secondary structural element to which each amino acid belongs is identified and for those that belong within loops, the position and length of the loop is presented.

derived from the NCBI database were aligned via the multiple sequence alignment tool. Generally there appears to be a high level of homology between luciferases however the candidate regions for targeted deletion in this work display variable levels of conservation.

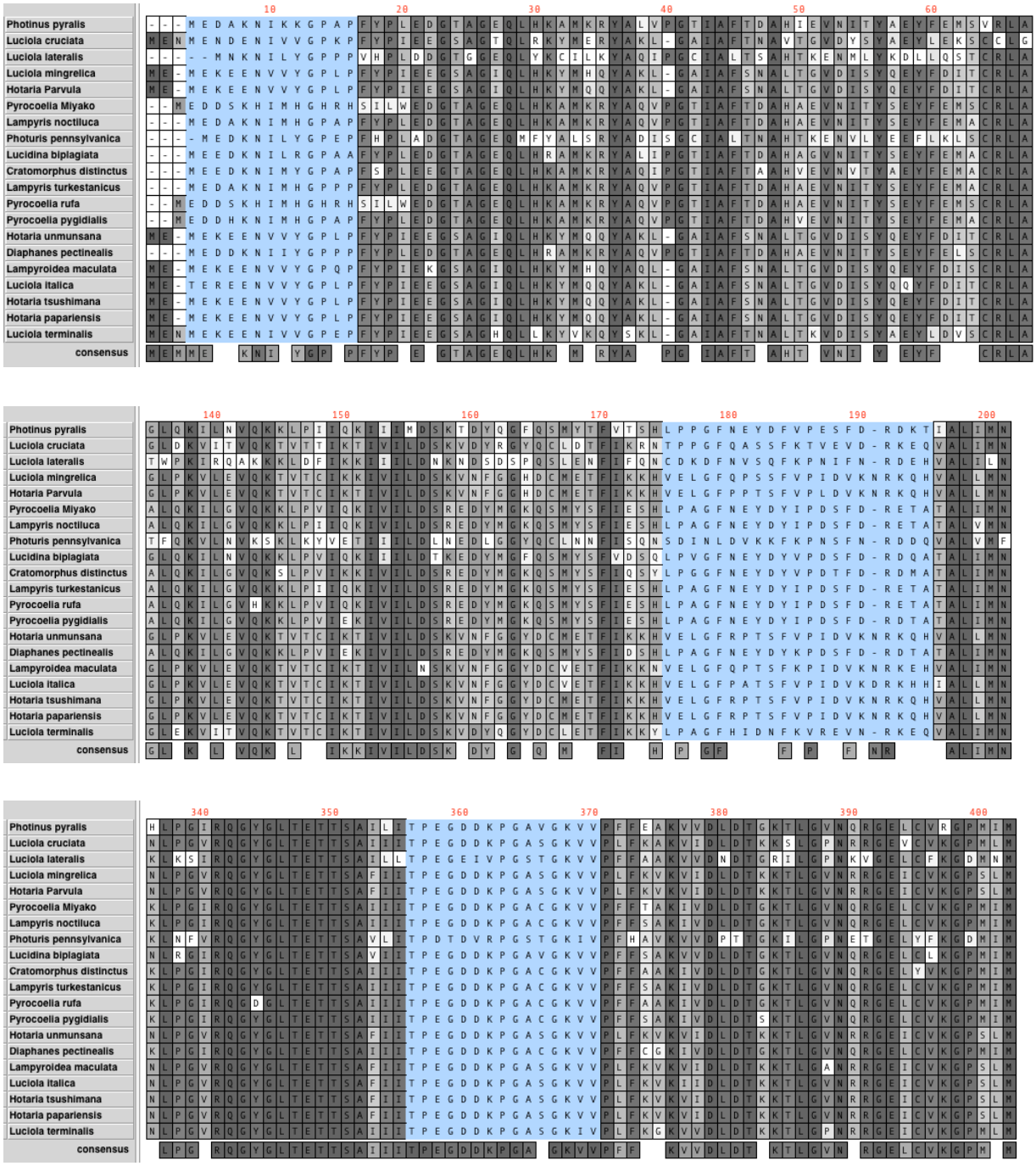
Standard analysis packages can determine conservation of individual amino acids within a protein however a method does not currently exist to quantify the overall conservation displayed by a particular loop (group of amino acids). To determine the level of conservation of a loop, a table was constructed to assess the number of amino acid residues within each loop grouped by percentage identity (Table 3.3.). Based on this data, a method was derived within this thesis to provide a conservation score.

The conservation score ranges from between 0-90 and acts as a measure to determine relative conservation of a loop whereby the greater the score, the greater the conservation of that loop. To derive this conservation score, the amino acids present within a loop are grouped by conservation identity of either >80%, 60-79%, 40-59%, 20-39% and 20% and the number of amino acids per group quantified, as Table 3.3. All data is converted to a fraction of the total amino acids present within that specific loop and these fractions are subsequently scaled utilising the mean % Conservation Identity of that group. Lastly, the conservation score is determined by the addition of each set of scaled fractions pertaining to a specific loop (Table 3.3.).

Of candidate regions, a conservation score of 56 and 60 was derived for the N- (M1-G10) and C-termini (K543-L550), respectively whilst the lowest score calculated pertained to L172-T191. On the other hand, conservation scores derived for T352-F368 and D520-L526 were 82.94 and 87.5, respectively. Therefore of the loops selected within this investigation, in order of conversation from most conserved to least conserved are T352-F368, D520-L526, K543-L550, M1-G10 and L172-T191.

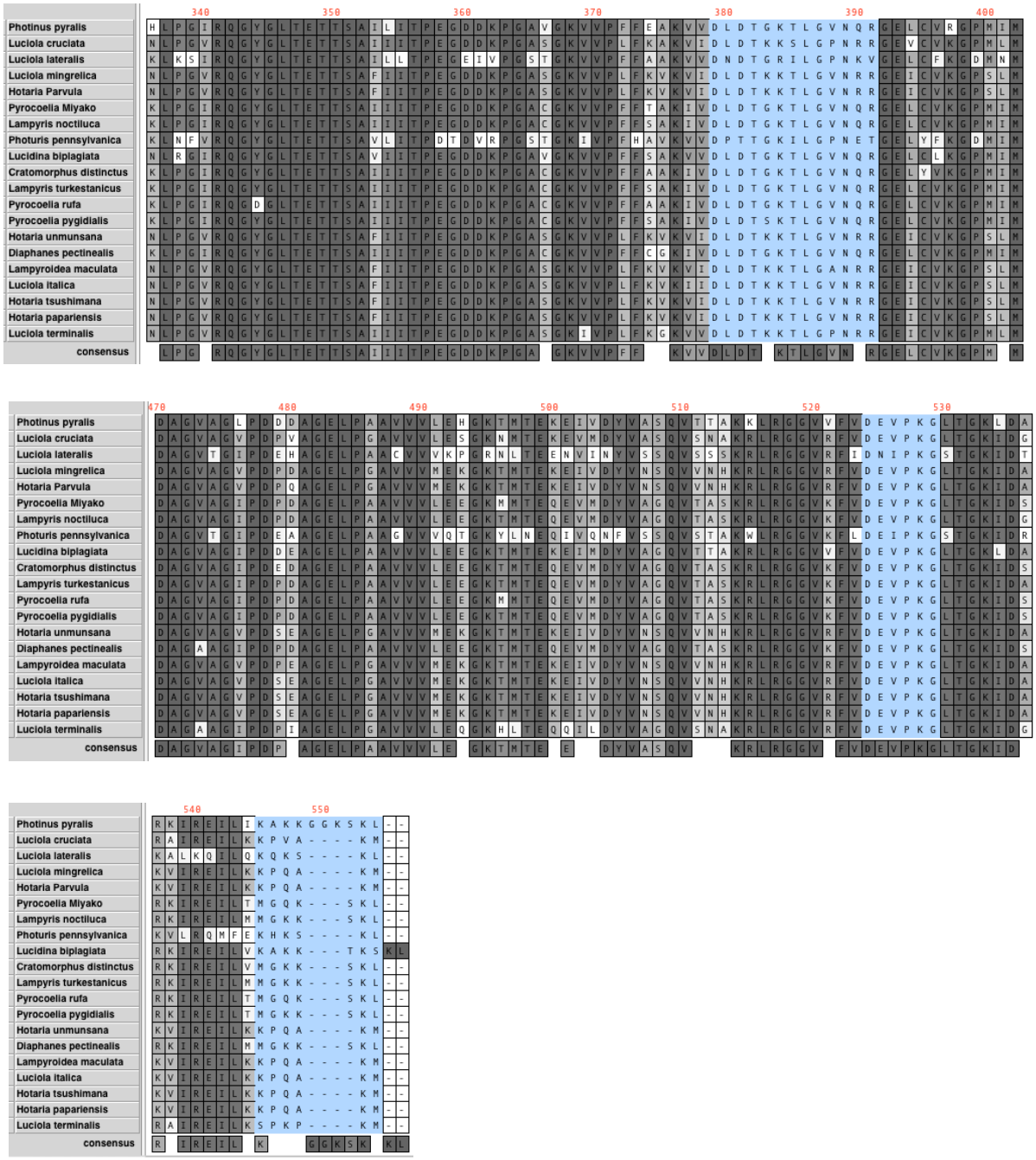
Of these candidate loops, 2 loops are better conserved (T352-F368 and D520-L526) whilst 3 loops show less conservation (M1-G10, L172- T191, K543-L550). It is desirable to select an additional loop with a conservation score similar to T352-F368 and D520-L526.

Figure 3.5. Multiple sequence alignment of Luciferases (Residues 1-403)



Sequence alignment of *P. pyralis* luciferase against selected beetle luciferases (Residues 1-403). A total of 20 homologues of beetle luciferase aligned on MacVector Inc (North Carolina, USA) to determine the level of conservation of target regions (highlighted blue)

Figure 3.6. Multiple sequence alignment of Luciferases (337 - 555)



Sequence alignment of *P. pyralis* luciferase against selected beetle luciferases (Residues 337 - 555). A total of 20 homologues of beetle luciferase aligned on MacVector Inc. to determine the level of conservation of target regions (highlighted blue).

Table 3.3. Conservation of Amino Acids identified from Multiple Sequence Alignment

Structure	**Number of Amino Acids Where Conservation Identity >80%**	**Number of Amino Acids Where Conservation Identity 60-79%**	**Number of Amino Acids Where Conservation Identity 40-59%**	**Number of Amino Acids Where Conservation Identity 20- 39%**	**Number of Amino Acids Where Conservation Identity <20%**	**Conservation Score**
M1-G10	4	1	2	0	3	56
L172- T191	5	1	2	0	13	35.71
T352-F368	15	0	1	0	1	82.94
D375-R387	9	0	0	0	2	75.45
D520-L526	7	1	0	0	0	87.5
K543-L550	5	0	0	0	3	60

Table showing the number of amino acids within each specific candidate loop, that have a conservation identity of either >80%, 60-79%, 40-59%, 20-39% and 20% as indicated by the consensus strand from multiple sequence alignment. A conservation score for each loop is highlighted. Conservation scores derived by converting all data in the table to a fraction of the total amino acids within that specific loop. These fractions are subsequently scaled utilising the mean % Conservation Identity and the conservation score was determined via the addition of each set of scaled fractions pertaining to a specific loop

Therefore, loop D375-R387, which exhibited a conservation score of 75.45 was selected as the last candidate loop for sequential single amino acid deletions (Figure 3.7).

3.3.6. Summary of Single Amino Acid Deletion Candidates

In summary of Sections 3.3.1. through to 3.3.6, loops that shall be targeted for single amino acid deletion include M1-G10, L172- T191, T352-F368, D375-R387, D520-L526, K543-L550 (Table 3.4. and Figure 3.7.). These loops were isolated for deletion based on analysis of the structure by a number of techniques including disorder prediction, residues occurring within 5Å of the active site and the identification of an omega loop amongst others.

3.3.7. One-Step Adapted Site-Directed Mutagenesis to Generate Single Amino Acid Deletions

Sequential single amino acid deletion within regions identified as targets (see Table 2.2.) was conducted by the careful design of mutagenic primers whereby a PCR reaction amplifies the whole plasmid whilst removing a trinucleotide corresponding to the selected deletion. Primers were designed such that the forward and reverse primer would not incorporate the targeted three nucleotides corresponding to a single amino acid in the final product as shown in Figure 3.8. For each targeted amino acid, the 5' end of both forward and reverse primers would amplify away from the trinucleotide to be deleted.

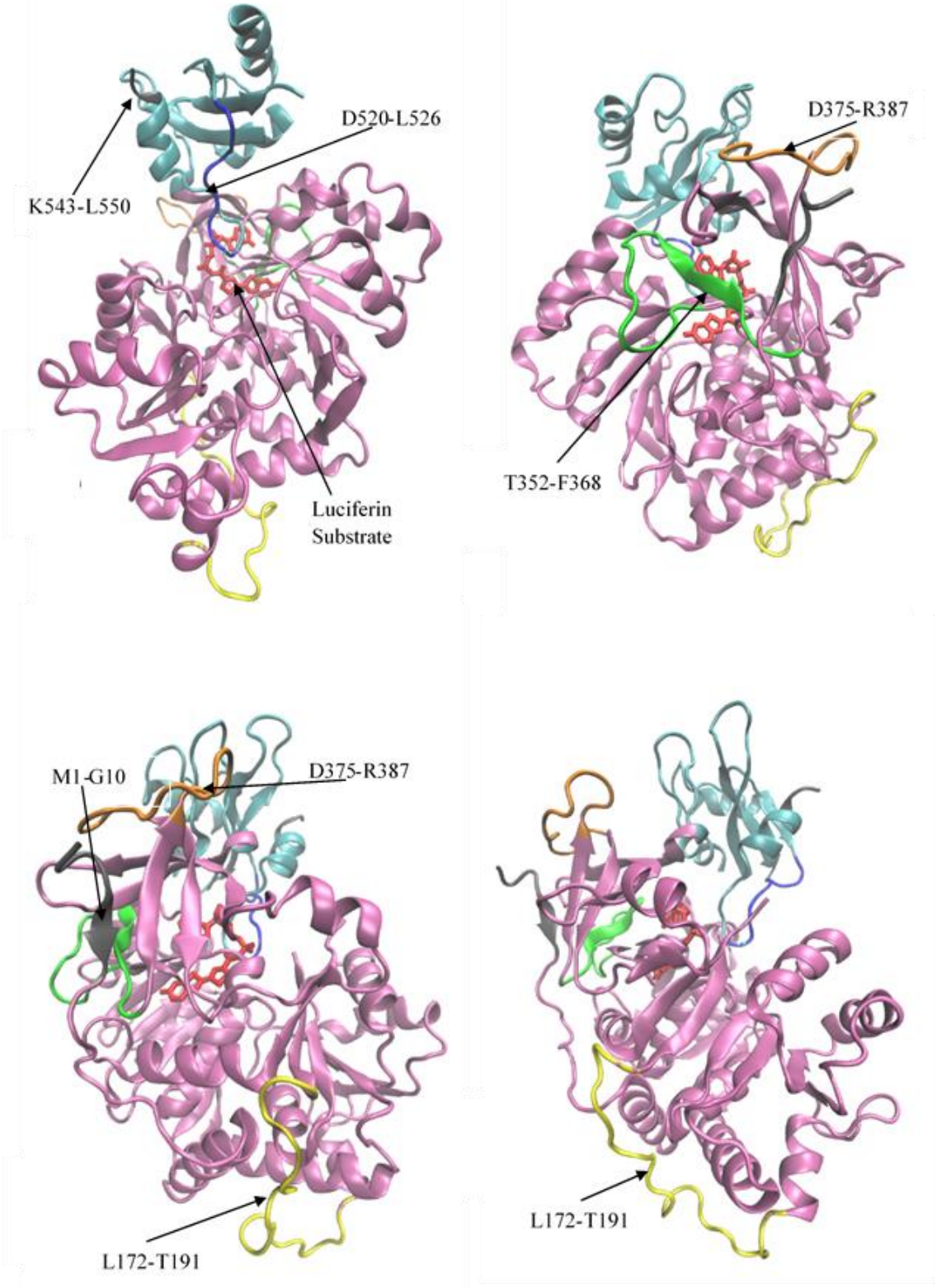
SDM reactions were transformed into BL21 (DE3) *E.coli* cells for screening and selection based on two criteria: brightness at RT and apparent resistance to thermal inactivation at 42^{o}C.

Table 3.4. Summary of Positions for Targeted Single Amino Acid Deletion

Region	Secondary Structure	Conservation Status	Additional Comment
M1-G10	N Terminal	56	
L172- T191	Loop	35.71	Disorded Loop
T352-F368	Loop	82.94	Omega Loop
D375-R387	Loop	75.45	
D520-L526	Loop	87.5	Within 5Å of Active Site
K543-L550	C Terminal	60	

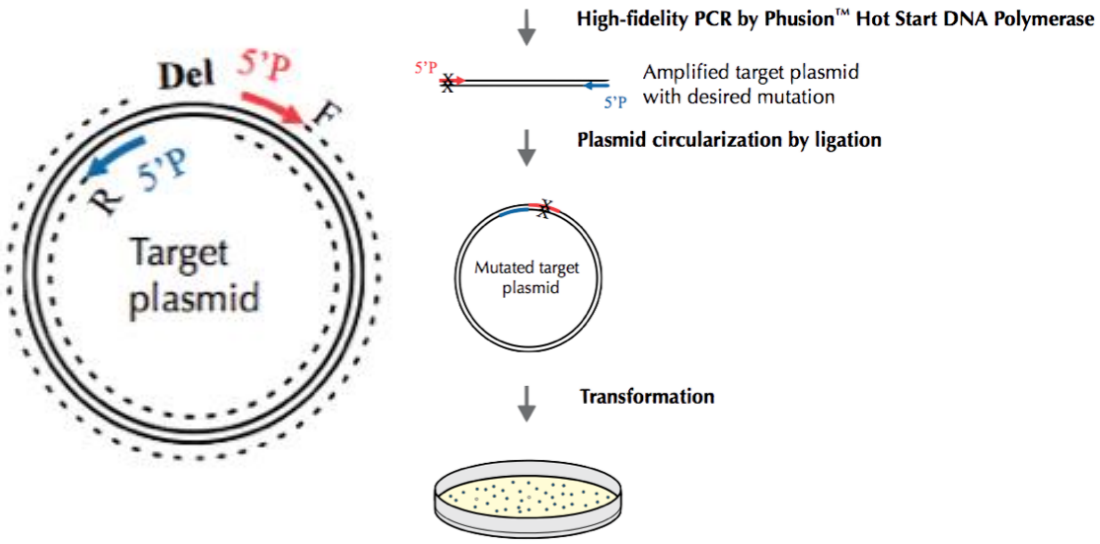
Summary the regions of the x11FLuc protein to be targeted for deletion and the corresponding secondary structure and the calculated conservation score.

Figure 3.7. Position of Loops Selected for Single Amino Acid Deletion



Rotated structure of substrate bound luciferase highlighting loops targeted for deletion. Loops M1-G10 (grey), L172- T191 (yellow), T352-F368 (green), D375-R387 (orange), D520-L526 (dark blue), K543-L550 (grey) are highlighted. Luciferin substrate (red). Image generated via VMD using PDB file ID: 4G36.

Figure 3.8. Illustration of site directed mutagenesis for the incorporation of single amino acids



Whole plasmid amplification of the pET16bx11FLuc for the introduction of single amino acid deletions. SDM carried out using primers as described by Table 2.2. used to generate single amino acid deletions at different positions throughout x11FLuc. Phusion polymerase was utilised in all experiments and *Dpn* I digestions of PCR products was carried out. Once mutated, the plasmid were directly transformed and screened in *E.coli* BL21 (DE3) cells (Figure donated by Dr. Dafydd Jones, Cardiff University).

3.3.8. Mutant Screening for Activity and Resistance to Thermal Inactivation

An effective screening method is key in any protein engineering strategy to identify possible candidates for purification and subsequent characterisation. The method initially trialed to identify potentially useful mutants in transformant colonies of *E. coli* was that reported by Wood and Deluca (1987) and improvement by Dr. Erica Law (PhD Thesis, University of Cambridge, 2004) to include a secondary screen.

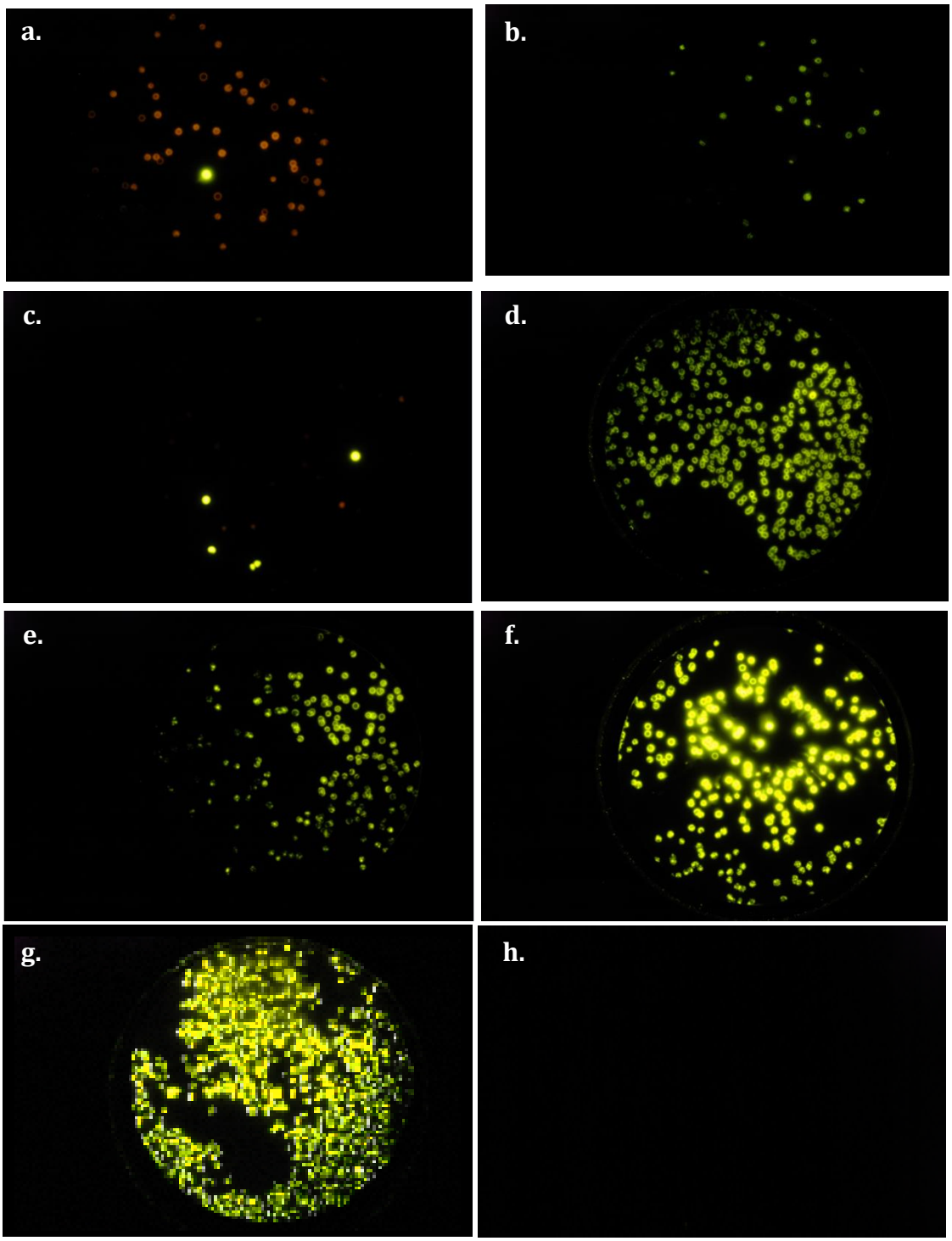
3.3.8.1. Primary Colony Screening and Sequencing Confirmation of Deletion Strategy

x11FLuc mutants generated from this site-directed deletion (SDD) strategy were initially screened for activity and sequenced in order answer key two initial questions; i) to identify whether bioluminescence was retained by the protein and ii) to confirm that the phenotype displayed was as a result of a single amino acid deletion.

All SDD reactions (with 62 primer sets) were transformed into *E.coli* and a primary screen was carried out. A representation of the primary screening conducted is illustrated in Figure 3.9. Ideally, all colonies present on any given transformant plate should be identical clones, containing the same single amino acid deletion and as a result should display identical phenotypes. However, multiple phenotypes were observed. Such variation is likely due either to incomplete *DpnI* digestion, allowing the appearance of the x11FLuc template or as a result of aberrant mutagenesis during PCR. Sequencing was conducted on any colonies that were representative of the different phenotypes identified on each plate to isolate those that contained the single amino acid deletion.

Of the 62 SDD reactions, sequencing confirmed that a total of 41 reactions had successfully incorporated the desired single amino acid deletion. Of the unsuccessful single amino acid deletion reactions, mutants had been generated which had promoted a frame shift by deletion or the incorporation of additional bases. In other cases, the PCR had been unsuccessful.

Figure 3.9. Representation of Primary Colony Screening



Representation of primary colony screening. Nitrocellulose membranes with *E.coli* BL21 (DE3) containing single amino acid deletions mutants of x11FLuc were induced for 3-4hours at RT with IPTG (1mM). a) ΔT378 b) ΔK544 c) ΔG379 d) ΔK543 e) ΔS548 f) ΔL550 g) x111Fluc (Positive Control) h) Negative Control. All plates were screened with 1mM LH_2 and then imaged on a Nikon D300 camera over an integration period of 30 seconds.

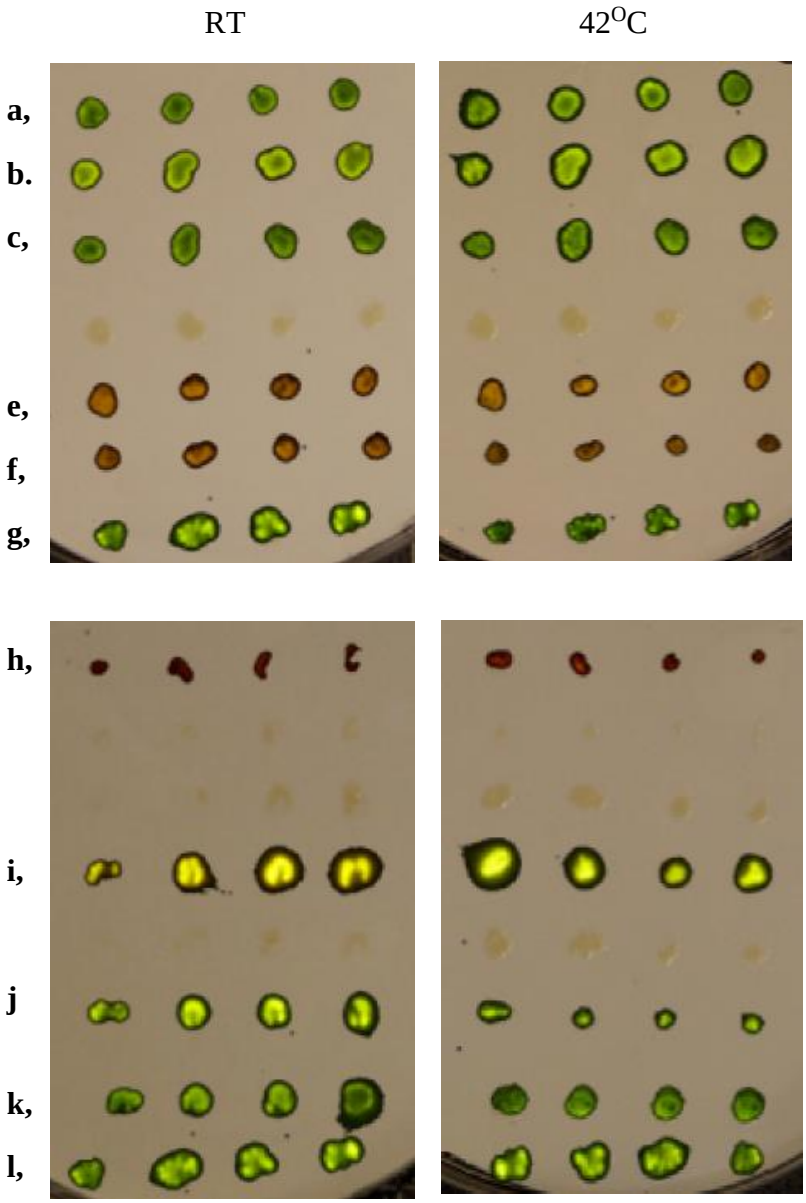
Sequencing additionally identified one mutant which had been generated incorporating two deletions (ΔF176 and ΔN177) which displayed bioluminescence activity, whereas a mutant which incorporated a single amino acid deletion in combination with a substitution, ΔP183/ E184D causes a loss of function. A focus of this investigation is to determine the overall tolerance of x11FLuc to single amino acid deletions. However, it is of interest that in the example where a double deletion, ΔF176/ ΔN177, has occurred, x11FLuc exhibits bioluminescent activity, while a loss of function has been has been observed due to ΔP183 in combination with a random substitution, E184D. Without further investigation it is difficult to attribute the loss of function of ΔP183/ E184D to either to the deletion, the substitution or a combination of both.

Of the 43 successful SDD reactions, 39 colonies displayed bioluminescence to varying degrees. In fact, only 2 variants, ΔV365 and ΔV366, exhibited a complete loss of activity. Both V365 and V366 are within 10Å of the ATP binding site and it is possible that the loss of function noted by ΔV365 and ΔV366 may be attributed to a loss in the ability for the substrate ATP to bind effectively to the protein.

3.3.8.2. Secondary Screening of Mutants

Significant variation was observed between plates during primary screening, which may be due to differences in incubation times or the lag time between spraying and collection of bioluminescence. Therefore, all 43 confirmed deletion mutants were subjected to a secondary screen (Figure 3.10). The purpose of implementing the secondary screen was to compare activity of the single amino acid deletion variants on one plate against the x11FLuc parental control in order to identify whether any mutants were likely to display increased bioluminescent activity. In addition to this, the secondary screen was implemented to assess whether the characteristic resistance to thermal inactivation displayed by the x11FLuc had been retained.

Figure 3.10. Representation of Secondary Screening of Mutants



Representation of Secondary Screening of Mutants. Nitrocellulose membranes with *E.coli* BL21 (DE3) containing single amino acid deletions mutants of x11FLuc were induced for 3-4hours at RT with IPTG (1mM). a-c) ΔD356, e-f) ΔK358, g) x11FLuc h) ΔT352, i) ΔP353 j) R354, k)G355, l) x11FLuc. All were firstly sprayed with 1mM LH_2 and then imaged on a Nikon D300 camera over an integration period of 30 seconds.

In order to conduct the secondary screen, single clones expressing the same single amino acid deletion were grown on two identical plates along with a control line of x11FLuc and two screens were carried out with and without incubation at 42ºC for 15-30 minutes. Following secondary screening, after incubation at 42ºC, of the 43 deletion mutants derived, all retained some activity. Even the double deletion mutant ΔF176/ ΔN177 displayed resistance to thermal inactivation at 42ºC by secondary screen, suggesting that the properties of x11FLuc were well maintained despite the location of the deletion.

This screening strategy is useful in answering basic questions in a library of mutants. However, the perception of brightness by an individual is subjective and an additional methodology would be advantageous for the quantification of bioluminescence during screening.

3.3.8.3. Image J Analysis for Colony Brightness at Elevated Temperatures

ImageJ is an image processing programme (National Institutes of Health, Bethesda, Maryland, USA) that may be utilised to overcome subjective observations and was utilised to quantify the overall light emitted from each colony in photographs of the secondary screens over a 30 second integration period. To do this, images were converted to 32-bit, manually defining the space in which the colony was to be analysed and the mean gray value, whose values are divided into 256 bins, were measured. The average bioluminescence was determined across each row of colonies expressing the same mutants and statistical tests applied (Figure 3.11.) (Table 3.5.)(Figure 3.12).

The spread of the data ranged from an average of 197.68 to 25.17, below the 256bins utilised within the software, indicating all values were not saturated. A trend was observed whereby deletions occurring in the N-terminal (M1-G10), displayed similar intensities to x11FLuc, with mean gray values of between 99-122 compared to 91 for x11FLuc. Deletions occurring within the C-terminal displayed values of between 172-174. Apparent resistance to thermal inactivation for both the N and C terminal were between 172-174, which appeared brighter than x11FLuc. With regards to the loop regions, deletions between L172- T191 displayed mean gray values of between 25-131, whereby 8 deletion mutations had reduced activity compared to x11FLuc whilst 3 were enhanced (Table 3.5, Figure 3.11). With regards to apparent resistance to thermal

inactivation at 42°C within this loop, surprisingly ΔP173, ΔD180, ΔF186 and ΔR188 had higher levels of activity compared to the results at RT. Deletions within loop T352-F368 displayed mean gray values of between 28-178. Eight deletion mutations had enhanced activity at RT compared to x11FLuc. However, only ΔT352 and ΔD356 in this loop promoted higher levels of activity at 42°C compared to RT. Deletions within loop D375-R387 displayed means gray values of between 8-125 and 1 deletion mutant had enhanced activity, whilst ΔG379, ΔN385 and ΔR387 promoted higher levels of activity at 42°C compared to RT. Lastly, D520-L526 displayed mean gray values of between 5-52, however, no deletion mutants had enhanced activity compared to x11FLuc, however again, all displayed higher levels of activity at 42°C compared to RT.

A trend was observed (Figure 3.11.) where deletions within the N-terminal displayed similar Image J results to x11FLuc, whilst C-terminal deletions promoted an enhancement of activity, suggesting that the N- terminal does not play a role in modulating activity. Interestingly, deletions occurring within the less conserved loop L172- T191 and the most stringently conserved loop, D520-L526 both resulted in reduced activity. This suggests that with regards to loops, the degree of conservation exhibited by a particular loop will not ultimately determine the overall tolerance of that deletion. Moreover, these effects indicate the importance of these loops for catalytic activity, suggesting residues present within these loops are implicated in the reaction/mechanism or indirectly co-ordinate residues that are involved. This shows the degree of conservation does not necessarily correlate to tolerance of a region to single amino acid deletion, for example, amongst the highest activity was displayed within T352-F368 (Ω loop). It has been considered that this Ω loop may be important for determining substrate specificity and the wide ranging activities displayed by deletions within this region are indicative that some modulation of the reaction is occurring.

With regards to Image J analysis of secondary screening for resistance to thermal inactivation, all deletion mutants retained activity after incubation at 42°C, however, the location of loop could not be used to predict the degree of resistance to thermal inactivation. N- and C-terminal deletions exhibited similar resistance to thermal inactivation as conferred by x11FLuc and also deletions within D520-L526 did not affect the phenotype.

Statistical analysis was conducted to determine the significance of the differences observed (Figure 3.12). A statistically significant difference between groups was found by a one-way ANOVA, $F(77, 154) = 42.67$, $p = 0.000$. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is not significant difference in activity between x11FLuc at RT and at 42^{o}C.

Figure 3.11. Image J Quantification of Bioluminescent Activity following Secondary Screening

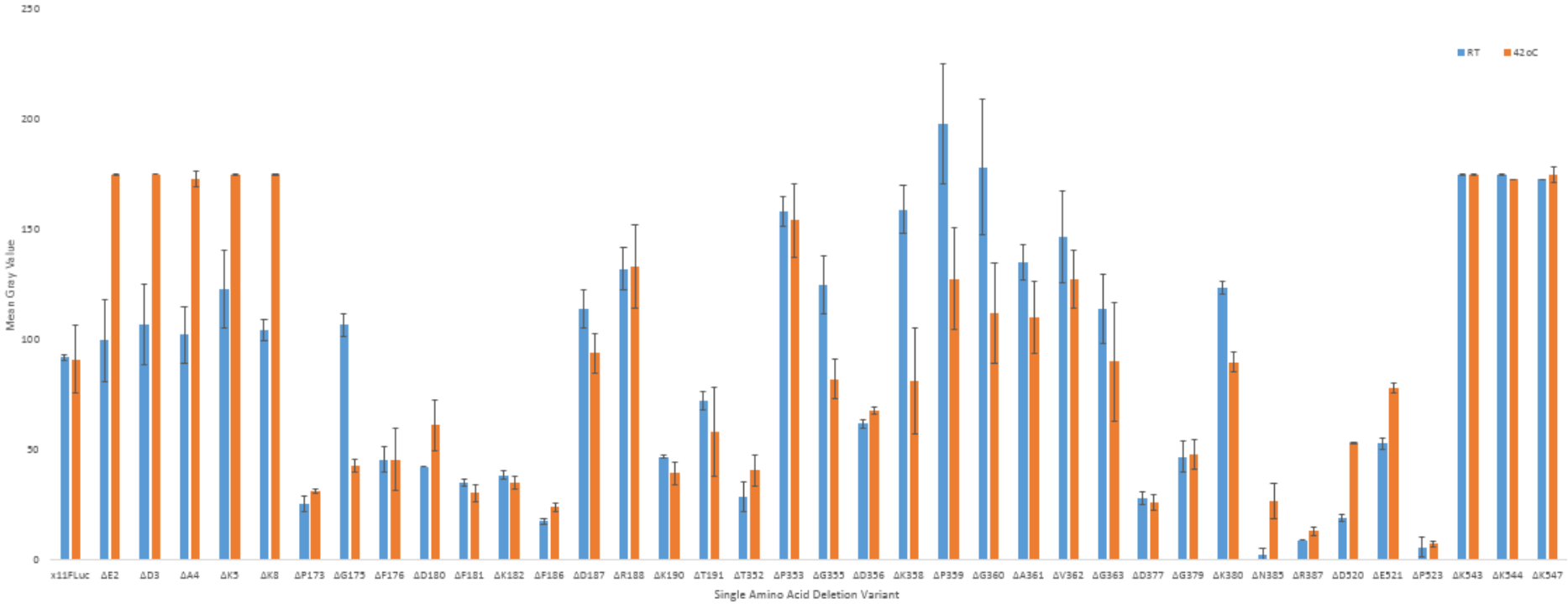


Image J analysis of colony brightness of mutants obtained from SDD. The data presented were obtained from secondary screening. Those bars in blue represent colony brightness at RT whilst those bars in red represent colony brightness at 42^{o}C. The average bioluminescence was determined by averaging of the mean gray values between each row of colonies expressing the same mutants and the standard error calculated (error bars).

Table 3.5. Image J Quantification of Bioluminescent Activity

Mutant	Average Bioluminescent Activity at RT	Standard Deviation at RT	Average Bioluminescent Activity at 42°C	Standard Deviation at 42°C
x11FLuc	91.71667	1.3811	90.95525	15.296
ΔE2	99.56	18.36374	174.4815	0.006414
ΔD3	106.5468	12.55756	174.6648	3.433884
ΔA4	101.9443	17.54107	172.5038	0.208003
ΔK5	122.6643	4.894652	174.533	0.05279
ΔK8	104.1933	3.390704	174.638	0.878785
ΔP173	25.1725	5.345311	31.0765	2.587806
ΔG175	106.4893	5.564744	42.648	14.12239
ΔF176	45.51025	0.195093	45.37525	11.54067
ΔD180	42.306	1.311436	61.203	3.765027
ΔF181	34.9655	2.13981	30.32425	2.757206
ΔK182	38.43375	1.481907	35.04675	1.907736
ΔF186	17.36075	8.794219	23.84375	8.859266
ΔD187	113.651	9.867305	93.71225	19.03901
ΔR188	131.9813	0.68552	133.0295	5.270331
ΔK190	46.58425	4.108691	39.22275	19.97117
ΔT191	72.00675	6.91147	58.095	6.9846
ΔT352	28.464	6.592046	40.43175	16.7816
ΔP353	157.7977	13.04921	153.8425	9.083142
ΔG355	124.8703	2.021349	81.9795	1.451428
ΔD356	61.6325	11.07943	67.67075	24.16931
ΔK358	158.928	27.14638	81.03125	23.00792
ΔP359	197.677	30.67849	127.2963	22.56761
ΔG360	178.0798	7.924617	111.7083	16.13466
ΔA361	134.7478	20.92089	109.8783	13.23135
ΔV362	146.5563	15.80108	127.1085	26.68634
ΔG363	113.5715	2.881687	89.769	3.549213
ΔD377	27.63625	7.249553	26.01675	6.57649
ΔG379	46.7025	2.739517	47.7225	4.600916
ΔK380	123.523	2.794327	89.55275	8.116116
ΔN385	2.09375	0.032987	26.81425	1.732864
ΔR387	8.91825	1.507288	12.8685	0.50384
ΔD520	19.10975	2.543419	52.79625	2.31786
ΔE521	52.59625	4.240126	78.0265	1.303664
ΔP523	5.78425	0.373989	7.34375	0.556983
ΔK543	174.4815	0.196984	174.6648	0.006414
ΔK544	174.6648	0.006414	172.5038	3.433884
ΔK547	172.5038	3.433884	174.533	0.208003

Image J analysis of bioluminescent activity within colonies following secondary screening. Bioluminescent activity observed within colonies of x11FLuc and deletion mutants at RT, and following incubation at 42°C. Bioluminescent activities measured by averaging of the mean gray values taken from 4 clone colonies.

Figure 3.12. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Variants at RT and 42ºC from Colony Screening

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	77	707873	9193.2	42.67	0.000
Error	154	33178	215.4		
Total	231	741051			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔK358 RT	3	183.7	A
ΔK544 RT	3	174.669	A B
ΔK543 42	3	174.669	A B
ΔD3 42	3	174.669	A B
ΔK8 42	3	174.622	A B
ΔK547 RT	3	174.547	A B
ΔK544 42	3	174.547	A B
ΔA4 42	3	174.547	A B
ΔK547 42	3	174.488	A B
ΔK5 42	3	174.488	A B
ΔK543 RT	3	174.420	A B
ΔE2 42	3	174.420	A B
ΔT352 42	3	163.21	A B C
ΔP359 RT	3	161.9	A B C
ΔK380 42	3	161.58	A B C
ΔT352 RT	2	157.5	A B C D E F
ΔD356 RT	3	153.98	A B C D
ΔA361 RT	3	143.1	A B C D E F G
ΔD187 42	3	140.4	A B C D E F G H
ΔD187 RT	2	135.9	A B C D E F G H I
ΔG360 RT	3	133.87	A B C D E F G H I
ΔA361 42	3	126.0	B C D E F G H I J
ΔG379 RT	3	122.22	C D E F G H I J
ΔP359 42	3	121.7	C D E F G H I J K
ΔP353 RT	3	119.7	C D E F G H I J K L
ΔK358 42	3	119.3	C D E F G H I J K L
ΔF186 RT	3	117.0	C D E F G H I J K L M
ΔG360 42	3	113.7	C D E F G H I J K L M
ΔK5 RT	3	112.60	C D E F G H I J K L M N
ΔG175 RT	3	108.13	D E F G H I J K L M N
ΔV362 RT	3	105.36	D E F G H I J K L M N
ΔK8 RT	3	102.27	F G H I J K L M N O
x11FLuc 42	**3**	**100.130**	**E F G H I J K L M N O**
ΔF186 42	3	96.70	G H I J K L M N O P
ΔD3 RT	3	95.6	G H I J K L M N O P
ΔA4 RT	3	95.2	G H I J K L M N O P Q
x11FLuc RT	**3**	**91.72**	**G H I J K L M N O P Q R**
ΔG379 42	3	91.66	G H I J K L M N O P Q R
ΔE2 RT	3	89.1	H I J K L M N O P Q R S
ΔD356 42	3	87.3	I J K L M N O P Q R S
ΔP353 42	3	84.56	I J K L M N O P Q R S T
ΔV362 42	3	82.6	I J K L M N O P Q R S T
ΔE521 42	3	78.03	J K L M N O P Q R S T U
ΔK380 RT	3	77.36	J K L M N O P Q R S T U V
ΔK190 RT	3	70.71	K L M N O P Q R S T U V W
ΔG355 42	3	68.522	L M N O P Q R S T U V W X
ΔK190 42	3	66.9	M N O P Q R S T U V W X Y
ΔG355 RT	3	62.11	N O P Q R S T U V W X Y Z
ΔF176+ΔN177 42	3	61.4	N O P Q R S T U V W X Y Z
ΔE521 RT	3	52.12	O P Q R S T U V W X Y Z AA
ΔD520 42	3	51.443	O P Q R S T U V W X Y Z AA
ΔD377 RT	3	51.01	O P Q R S T U V W X Y Z AA
ΔF176 42	3	48.3	P Q R S T U V W X Y Z AA
ΔF176 RT	3	46.78	P Q R S T U V W X Y Z AA
ΔR188 RT	3	46.307	P Q R S T U V W X Y Z AA

ΔD377	42	3	46.04	P Q R S T U V W X Y Z AA
ΔT191	42	3	43.87	Q R S T U V W X Y Z AA
ΔG175	42	3	42.81	R S T U V W X Y Z AA
ΔR188	42	3	42.39	R S T U V W X Y Z AA
ΔF176+ΔN177	RT	3	42.19	R S T U V W X Y Z AA
ΔF181	RT	3	39.619	S T U V W X Y Z AA
ΔF181	42	3	35.63	T U V W X Y Z AA
ΔD180	RT	3	34.40	T U V W X Y Z AA
ΔP173	42	3	30.80	U V W X Y Z AA
ΔD180	42	3	29.29	U V W X Y Z AA
ΔG363	42	3	26.81	U V W X Y Z AA
ΔN385	42	3	26.39	V W X Y Z AA
ΔG363	RT	3	26.31	V W X Y Z AA
ΔP173	RT	3	25.81	W X Y Z AA
ΔT191	RT	3	24.68	W X Y Z AA
ΔK182	42	3	24.51	W X Y Z AA
ΔD520	RT	3	18.48	X Y Z AA
ΔK182	RT	3	16.500	Y Z AA
ΔR387	42	3	13.149	Z AA
ΔR387	RT	3	9.802	AA
ΔP523	42	3	7.109	AA
ΔP523	RT	3	5.718	AA
ΔN385	RT	3	2.0833	AA

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from Image J analysis from bioluminescent colonies. A statistically significant difference between groups was determined by a one-way ANOVA, F(77, 154) = 42.67, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. .x11FLuc is highlighted in bold. RT= Variant at Room Temperature. 42= Variant at 42 ^{o}C. N=3.

3.3.8.4. Spectral Shifts Resulting from Single Amino Acid Deletions.

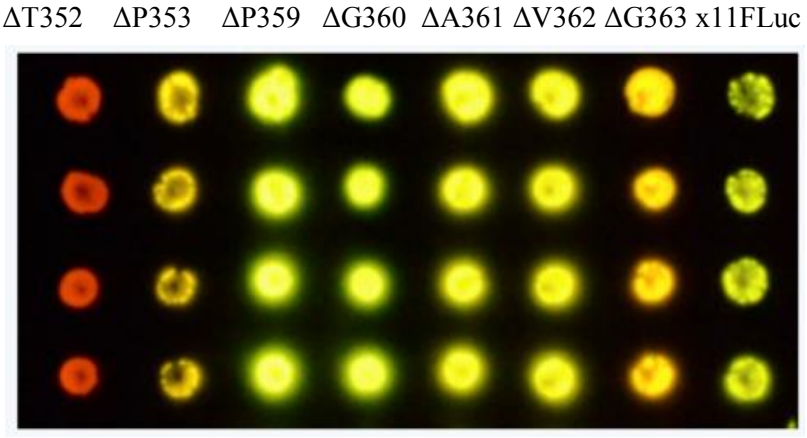
It was observed through *in vivo* colony screening that single amino acid deletions can alter the colour of bioluminescent emission. Five replicate colonies expressing deletion mutants within the Ω loop of the x11FLuc were screened (Figure 3.13-3.14.). It is intriguing that the deletions in close proximity to the ends of the N- and C-termini of the Ω loop, such as mutants ΔT352, ΔP353 and ΔG363 exhibit larger red shifts than deletions closer to the centre (e.g. ΔP359, ΔG360, ΔA361 and ΔV362), which exhibit emission colour similar to that of x11FLuc.

To study the bioluminescent spectra of the deletion variants, proteins were overexpressed and crude cell lysates were utilised to obtain bioluminescent spectra (Table 3.6.). Of a total of 43 mutants, a majority of deletion mutants (38) displayed a similar bioluminescent emission peak to the x11FLuc (λmax = 555nm). However, a red shift of 50nm was observed for mutant ΔT352 (λmax = 605nm) whilst ΔG363 displayed a smaller shift of 16nm (λmax = 571nm) (Figure 3.13). In addition to this, deletions within the loop D520-L526 promoted similar red shifts (λmax = 571nm). The majority of mutants exhibited a bioluminescence spectral half-bandwidth (full width at half maximum, FWHM) of between 51-61nm, similar to that of x11FLuc (60nm) whilst deletions within T352-F368 exhibited FWHM of between 38-74nm.

It is interesting that the majority of deletions throughout the structure of x11FLuc have not resulted in large alterations in the bioluminescence spectra of x11FLuc, considering that single amino acid substitution (S284T) results in a ca. 50nm red shift (Branchini *et al.*, 2005). However, one particular region, the Ω loop results in large alterations in λmax and FWHM. Previous *in vivo* colony screening of deletion mutants within this region showed a number of spectral changes. Based on this data, ΔP353, ΔP359, ΔG360, ΔA361 and ΔV362 primarily displayed changes in FWHM. ΔT352 displayed a prominent shift in λmax, while ΔG363 appears to be a combination of both.

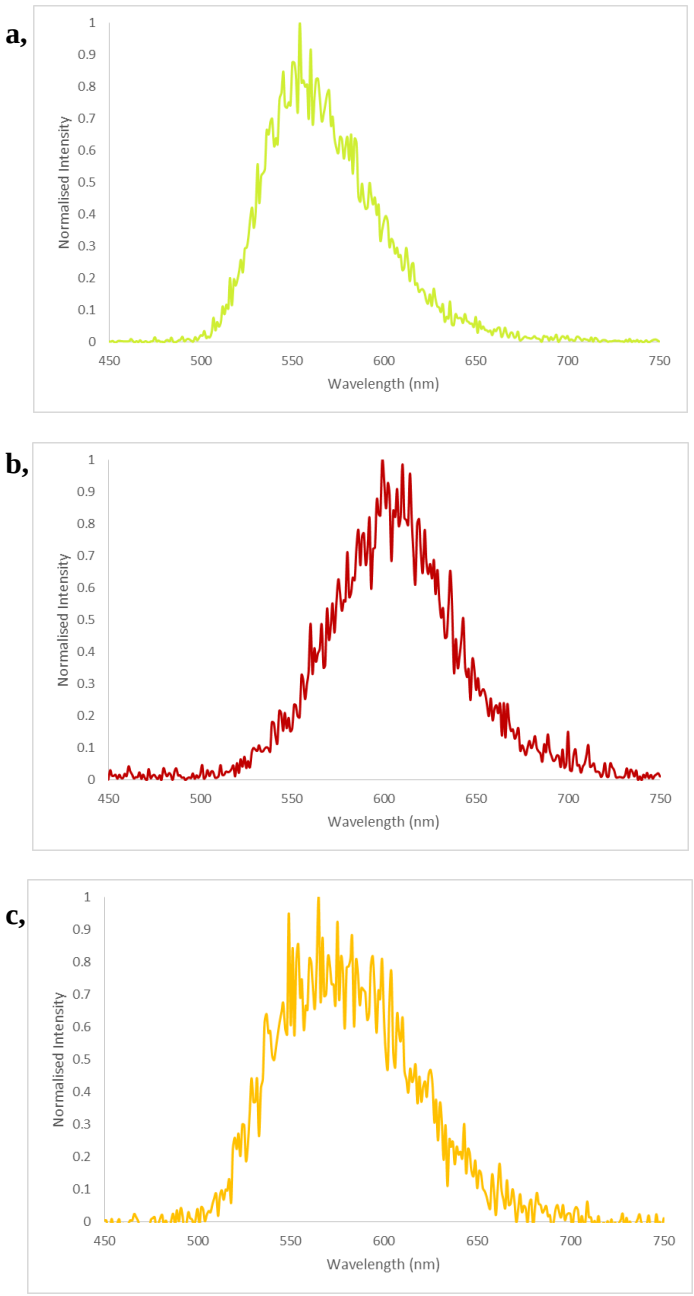
As suggested by colour theory, the differences noted here may be attributed to a number of factors. FWHM provides information on the distribution of the population of emitting species. However, with regards to ΔT352 whereby no such alteration in FWHM has occurred, shifts may be due to micro-environmental changes (see Chapter 1).

Figure 3.13. Colour shifted single amino acid deletion mutants



Colour-shifted single amino acid deletions mutants. Nitrocellulose membrane of *E.coli* BL21 (DE3) containing single amino acid deletions within x11FLuc (ΔT352, ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363) compared to native x11FLuc, induced for 210 min at RT with IPTG (1mM). All were sprayed three times with 1mM LH_2 in citrate buffer (pH 5.0) and imaged on a Nikon D300 camera over an integration period of 30 seconds.

Figure 3.14. Representation of Emission Spectra from Crude Cell Lysates of Omega loop mutants



Representation of Emission Spectra from Crude Cell Lysates. Crude cell lysates of *E.coli* BL21 containing a) x11Fluc b) ΔT352 c) ΔG363 induced for 6 hours at RT with IPTG (1mM). Spectral emission measured within Cary Eclipse fluorimeter following injection of 500uM LH_2 and 1mM ATP.

3.4. Further Discussion

The overarching aim of this investigation was to address whether single amino acid deletions could be tolerated by x11FLuc a thermostable and pH-tolerant mutant of *Ppy* WTFLuc. In order to answer this question, a targeted approach (SDD) was utilised to incorporate the single amino acid deletions as an alternative to substitution or random deletion mutagenic methods, such that specific loops are targeted for deletion and this provides a rational survey of the effects of deletions in loop structures.

Analysis of the secondary and crystal structures indicated that N- and C-terminals and an internal loop were good candidates for deletion due to high relative disorder. In addition to this, another and previously overlooked structure within the protein was identified as an Ω loop, a well characterised structure with links to protein specificity and activity in other proteins (Fetrow *et al.*, 1995). Furthermore, loops were surveyed within 5Å of the active site to identify if deletions would be tolerated when in close proximity to the catalytic site as such M1-G10, L172- T191, T352-F368, D375-R387, D520-L526, K543-L550 were selected to undergo SDD.

The screening and selection of mutants resulting from any mutagenesis strategy is a crucial step in identifying mutants of desired characteristics for further study. A convenient colony screening strategy has been utilised for the initial identification of mutants displaying bioluminescence activity. The protocol utilised was a method described by Wood and DeLuca (1987) which was useful for high throughput identification of tolerant mutations, however there were some limitations and biases within using the technique.

Firstly, whilst the *in vivo* colony screening method was useful for determining whether deletions were either tolerated or not tolerated, the subjective nature of the estimation of brightness by the human eye was considered problematic when aiming to quantify brightness in a large number of mutants. Image J analysis was utilised to overcome such problems however some bias remained. For example, often there is not 100% efficiency in the lifting of colonies using a nitrocellulose membrane. Secondly, another common problem involves obtaining an even spray of the LH_2 substrate and as a result colonies towards the periphery of the plate may be less exposed to the substrate compared to those cells within the centre of the plate and they may appear dimmer. Practically, there

is a lag time between the spraying and the taking of the photograph and this can be variable. Decay in the luminescent signal between spraying and image acquisition can result in differences in the apparent brightness of colonies.

Variations in brightness may also be attributed to the expression or the stability of the luciferase itself within the *E.coli,* where the expression pattern of x11FLuc may be different to those of the mutants. For example, we may consider that single amino acid deletions can impact upon the folding efficiency of an enzymes (Arpino *et al.,* 2014) changing the landscape of folding in favour of the production of soluble protein or perhaps a more stable protein appears more active after induction at RT (Jathoul *et al*., 2012).

At present, utilising deletional mutagenesis 43 deletion mutants were generated (Figure 3.15) and 41 retained bioluminescence activity, whilst 2 mutants were knock-outs. This is remarkable considering these mutants contain deletions in different loops connecting important areas throughout their structure. As hypothesised, prior to this survey (based on previous work with GFP), luciferase being a more globular and lesser highly structured protein is able to tolerate single amino acid deletions both within regions close to the active site and also remarkably within sites that are highly conserved. The loss of function within the mutants ΔV365 and ΔV366 indicates that these residues play a vital role in bioluminescence either directly or indirectly via H-bond networks with other residues or water molecules (Nakatsu *et al.,* 2006).

Further characterisation was conducted to assess the tolerance of x11FLuc to deletions in terms of characteristics such as thermostability and emission colour (Table 3.6.). Investigation showed that x11FLuc retained bioluminescence up to 42^{o}C despite the location of the single amino acid deletion. One key difference in mutants compared to x11FLuc was in reference to their bioluminescence emission □max and FWHM. Both specific and non-specific effects on the emitter were seen using crude cell lysates, as indicated by shifts with and without a change in spectral FWHM.

Despite improvements in this traditional *in vivo* colony screening technique, there are a number of drawbacks which lead to results which are difficult to interpret to select mutants for further characterisation from a large pool. To further improve the screening

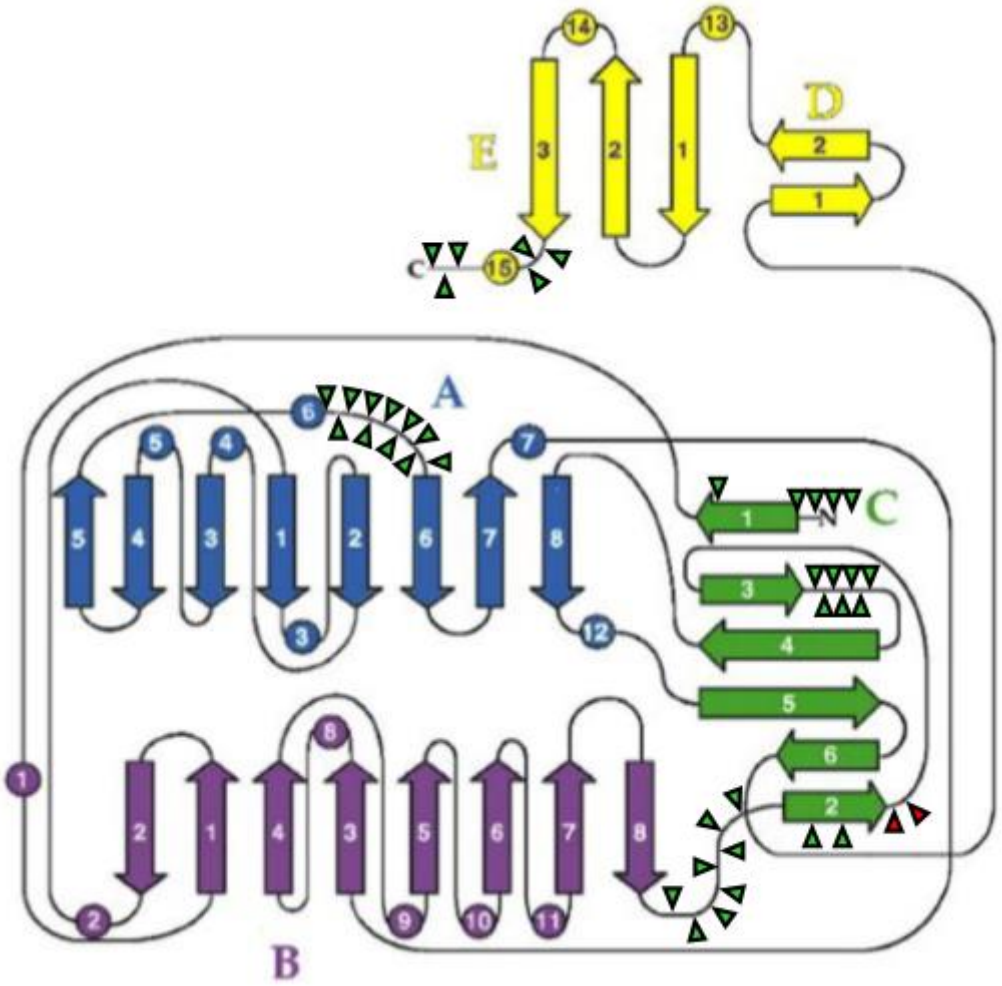
method optimisation to accelerate the identification of useful mutants. In particular, a method able to quantify bioluminescent activity would be desirable.

Table 3.6. Summary of Bioluminescence Properties of Deletion Mutants from Colony Screening and Crude Cell Lysates

Variant	**Structure**	**Bioluminescence Activity**	**Apparent Resistance to Thermal Inactivation (42ºC)**	**Bioluminescence λmax (nm)**	**FWHM (nm)**
x11FLuc	Control	✓	✓	555	60.97
E2	N Terminal	✓	✓	ND	ND
D3	N Terminal	✓	✓	ND	ND
A4	N Terminal	✓	✓	ND	ND
K5	N Terminal	✓	✓	ND	ND
K8	N Terminal	✓	✓	ND	ND
P173	L172- T191	✓	✓	555	54
G175	L172- T191	✓	✓	555	59.91
F176	L172- T191	✓	✓	555	61.05
F176+N177	L172- T191	✓	✓	555	55.99
D180	L172- T191	✓	✓	555	58.06
F181	L172- T191	✓	✓	555	58.04
K182	L172- T191	✓	✓	555	61.02
P183+E184D	L172- T191	×	×	×	×
F186	L172- T191	✓	✓	555	56.99
D187	L172- T191	✓	✓	555	51
R188	L172- T191	✓	✓	555	54.87
K190	L172- T191	✓	✓	555	52.94
T191	L172- T191	✓	✓	555	56.9
T352	T352-F368	✓	✓	605	62.04
P353	T352-F368	✓	✓	555	48.04
G355	T352-F368	✓	✓	555	57.97
D356	T352-F368	✓	✓	555	38.05
K358	T352-F368	✓	✓	555	54.95
P359	T352-F368	✓	✓	555	56.9
G360	T352-F368	✓	✓	555	59.98
A361	T352-F368	✓	✓	555	64
V362	T352-F368	✓	✓	555	52.93
G363	T352-F368	✓	✓	571	73.9
V365	T352-F368	×	×	×	×
V366	T352-F368	×	×	×	×
L376	D375-R387	✓	✓	555	ND
D377	D375-R387	✓	✓	555	ND
G379	D375-R387	✓	✓	555	ND
K380	D375-R387	✓	✓	555	ND
V384	D375-R387	✓	✓	555	ND
N385	D375-R387	✓	✓	555	ND
R387	D375-R387	✓	✓	555	ND
D520	Proximal to	✓	✓	571	ND
E521	Proximal to	✓	✓	572	ND
P523	Proximal to	✓	✓	571	ND
K543	C Terminal	✓	✓	555	ND
K544	C Terminal	✓	✓	555	ND
K547	C Terminal	✓	✓	555	ND

A combination of data from colony screens and crude cell lysates to illustrate properties of x11FLuc deletion mutants. Bioluminescence spectra □max and FWHM are approximate values determined from crude cell lysates. ND: not determined.

Figure 3.15. Position of Successful Deletions in the Secondary Structure of x11FLuc



Position of Amino Acid Deletions made within x11FLuc (adapted from Conti *et al*., 1996). The secondary structure is shown, coloured by domains. Those single amino acid deletions retaining bioluminescence are highlighted via green triangles whilst those single amino acids deletions where bioluminescence is lost is marked via red triangle.

3.5. Conclusion

A SDD approach was used to mutate x11FLuc in specific positions chosen both for likely tolerance to single amino acid deletion and to investigate particular regions of interest in the luciferase protein. Analysis of the secondary and crystal structures indicated that N- and C-termini and an internal loop were good candidates for deletion due to high relative disorder. In addition to this, another and previously overlooked structure within the protein was identified as an Ω loop, a well characterised structure with links to protein specificity and activity in other proteins. Furthermore, loops were surveyed within 5Å of the active site to identify if deletions would be tolerated when in close proximity to the active site.

This targeted approach and subsequent screening of the mutants for brightness and apparent thermostability revealed that, in the main, mutants retained bioluminescent activity and the thermostable properties of x11FLuc parent up to 42^{o}C. In addition, it was revealed that the level of the conservation of the loops did not dictate tolerance to deletion.

Whilst the traditional *in vivo* colony screening technique is useful in determining the tolerance of the amino acids to deletion, there are drawbacks which lead to results which are difficult to interpret to select mutants for further characterisation from a large pool. To further improve the screening method optimisation is required to overcome some of these the drawbacks and accelerate the identification of useful mutants. In particular, a method able to quantify bioluminescent activity would be desirable and that is the focus of the next Chapter.

Chapter 4

Optimisation of Screening Strategies to Identify Useful x11 FLuc Deletion Mutants

4.1 Chapter Summary

Whilst conventional methodologies of colony screening activity in a large library of mutant luciferases are useful for determining certain characteristics such as apparent resistance to thermal inactivation, they are less suitable for the determination of other desirable characteristics, such as changes in kinetics or the specific activity of the enzyme, increased tolerance to pH and increased resistance to inhibition. Conventionally, such characteristics may not be determined until after proteins have been overexpressed, purified and characterised, which is both costly and time consuming. Therefore, methods to more accurately quantify bioluminescence under a variety of assay conditions prior to purification are desirable. Therefore 96-well format luminometric cell lysate assays were developed to refine the selection of the x11FLuc deletion library. Utilisation of this technique provided robust results with which it was possible to select mutants for purification and analysis using a wide range of selection criteria, such as resistance to inhibition by a number of factors. The Ω loop of x11FLuc was also confirmed to be an important structural feature, whereby single amino acid deletions cause wide-ranging phenotypic effects, such as alteration of bioluminescence kinetics, bioluminescence spectra, resistances to thermal inactivation and brightness.

4.2. Introduction

Colony screening of recombinant FLucs was first described by Wood and Deluca (1987). However, this method proved to be limited for the selection of Lucs with different beneficial characteristics from the deletion mutant library generated in Chapter 3. For example, desirable characteristics for ATP monitoring include enhanced

activities, kinetics, resistance to inhibitors, such as dehydroluciferin, spectral stability, and resistance to thermal inactivation or changes in pH.

FLuc reaction kinetics describe the rates of the bioluminescence reaction. The Michaelis constant (K_M) provides an indication of substrate affinity of enzymes, whilst k_{cat} is a measure of the turnover of FLuc and much protein engineering has been used to modulate these characteristics. The LH_2 substrate is a key cost in technologies that utilise FLuc therefore mutants identified with a reduced K_M for LH_2 are desirable within industry, in addition, kinetically altered enzymes provide insight into the mechanism of FLuc which is desirable within academia. Lucs able to resist inhibitors are also highly desirable, for example, FLuc is strongly inhibited by dehydroluciferin (L), which is an oxidation product of LH_2 and this can impact on the reproducibility of assays, therefore, mutants which resist such inhibitors may be of benefit to diagnostic applications of FLuc.

Characterisation of FLucs following purification provides the most meaningful analysis of kinetic and inhibition parameters, however this can limit the ability to quickly select multiple useful phenotypes from large mutant libraries. Furthermore, less bright mutants which have other useful characteristics may not be selected. Therefore, it was hypothesised that a screening strategy could be optimised which can better determine a number of Fluc characteristics prior to enzyme purification. The development of a new screening technique would allow for characterisation of a greater number of novel deletion mutants and provide more information about mutants that could be overlooked by conventional colony screening methods.

The aim of this Chapter is to adapt the conventional library screening methods into high throughput 96-well format assays for activity and resistance to inhibition or inactivation in conditions relevant to in vitro and field diagnostics, such as at physiological pH and at RT. Furthermore, this method will be used to better characterise all 43 deletion mutants of x11FLuc generated in Chapter 3 to help selection of bright, stable and less inhibited mutants for in vitro diagnostics applications.

4.3 Results and Discussion

4.3.1 Construction of plasmid pET16b-x2

The x2 Fluc (White et al., 2001) gene was cloned into the pET16b vector for to compare the properties of this mutant to x11 Fluc and its deletion mutants using 96-well format assays (see Chapter 2). This construct was named pET16b-x2.

4.3.2 Construction of a 96-Well Format Screening Strategy in Different Assay Conditions

In Chapter 3, 43 deletion mutants were screened to identify whether luciferase tolerated single amino acid deletions and whether the resistance to thermal inactivation typically displayed by x11FLuc had been retained. The results highlighted that in all bar two of the deletion mutants activity had been retained, however, it is not feasible to purify and characterise all 41 mutants, as such, a method able to better select desirable properties form this library was developed as an alternative to colony screening for better quantification of bioluminescence, to indicate low K_M for LH_2, and determine resistance to inhibitors.

To test whether this strategy would be robust enough to identify useful mutants, controls (WTFLuc, x2FLuc and x11FLuc) were overexpressed in *E. coli*, normalised to their $O.D._{600}$, lysed and resuspended in TEM buffer before assaying for bioluminescence by injecting luminometry. The assays were performed with a settle time of 2000ms and followed by an integration period of 5000ms as to capture steady state bioluminescence following the flash. Three conditions were utilised for testing, saturating substrate conditions, non-saturating LH_2 conditions and a PPi assay.

Results presented that under saturating substrate conditions WTFLuc, x2FLuc and x11FLuc exhibited activities of 12359226 RLU, 47009667 RLU and 29399000 RLU respectively, whilst non-saturating LH_2 conditions exhibited activities of 86943 RLU (0.7%), 1296500 RLU (2.75%) and 626920 RLU (2.13%), respectively. The PPi assay

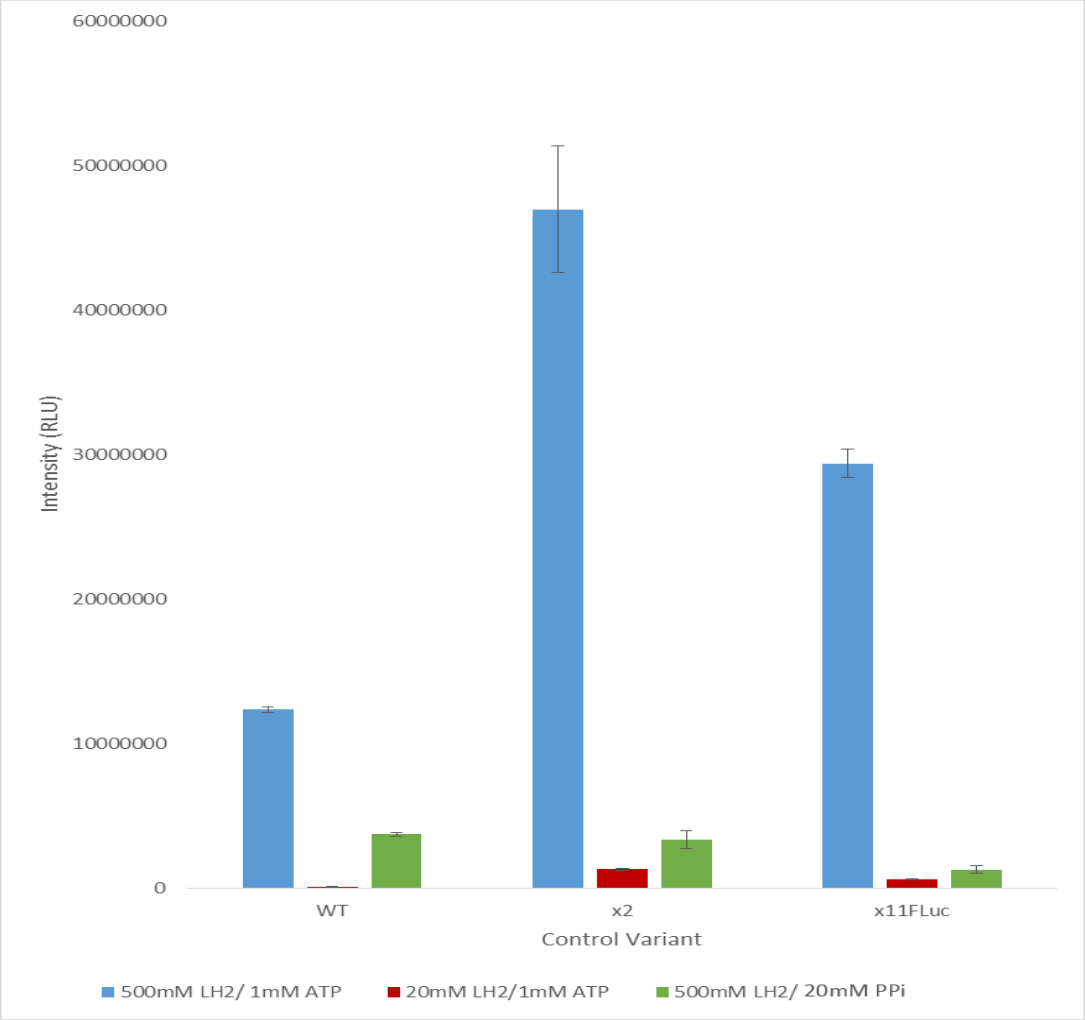
exhibited activities of 3734667 RLU (30.2%), 3351467 RLU (7.12%) and 1291270 RLU (4.39%), respectively (Figure 4.1.) (Table 4.1.).

Previous investigations note that at RT x2FLuc displays greater activity compared to WTFLuc and x11FLuc, a result which correlates to what was seen using the cell lysate assay (Figure 4.1). In addition to this, it is known that PPi is an inhibitor of the reaction and the results were concurrent with this, showing a reduction in activity within all 3 controls. Lastly, a screen was conducted to test whether the screening methodology would be able to discriminate differences in K_M. Of the controls, within the literature, WTFLuc has the highest K_M for Luciferin whilst x11FLuc has the lowest. In substrate limiting conditions, WT presented the lowest activity suggesting that during large scale screening those mutants with lower activity given such conditions are more likely to be characterised with a higher K_M.

Statistical analysis was conducted to determine the significance of the differences observed (Figure 4.2). A statistically significant difference between WT, x2 and x11FLuc was found by a one-way ANOVA, $F(2,6) = 47.35$, $p = 0.000$. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is a significant difference in activity of these controls under saturating conditions.

Further analysis was conducted to determine the significance of the difference between WT, x2 and x11FLuc under given conditions. Under limiting substrate conditions, a statistically significant difference between all variants was found by a one-way ANOVA, $F(2,5) = 179.81$, $p = 0.000$ and confirmed by the Tukey method (Figure 4.3) whilst under inclusion of 2mM PPi, a one-way ANOVA, $F(2,6) = 12.39$, $p = 0.007$ and Tukey test identified a non significant difference between WT and x2 and a significant difference between x11FLuc compared to WT and x2 (Figure 4.4).

Figure 4.1. Bioluminescence activity of Control Variants under 500µM LH_2/ 1mM ATP, 20µM LH_2/1mM ATP and 500µM LH_2/ 1mM ATP/2mM PPi



Bioluminescence activity displayed by controls under varying conditions. Crude cell lysates, normalised by OD_{600} of *E.coli BL21* containing WTFLuc, x2FLuc and x11FLuc, induced for 6 hours at RT with 1mM IPTG. Bioluminescence activity measured within Fluoroskan Ascent luminometer following injection of 500□M LH_2 with 1mM ATP, 500□M LH_2 with 1mM ATP with 2mM PPi and 20□M LH_2 with 1mM ATP.

Table 4.1. Bioluminescence Activity of Control Variants under 500µM LH_2/ 1mM ATP, 20µM LH_2/1mM ATP and 500µM LH_2/ 1mM ATP/500mM PPi

Variant	Mean 500µM LH_2/ 1mM ATP	St.Dev 500µM LH2/ 1mM ATP	Mean 20µM LH_2/1mM ATP	St. Dev 20µM LH_2/1mM ATP	Mean 500µM LH_2/ 1mM ATP/2mM PPi	St. Dev 500µM LH2/ 1mM ATP/2mM PPi
WT	12359226	215136.7	86943 (0.7%)	13254.09	3734667 (30.21%)	136345
x2	47009667	4351651	1296500 (2.75%)	66820.14	3351467 (7.12%)	616594.5
x11FLuc	29399000	994018.8	626920 (2.13%)	26633.58	1291270 (4.39%)	290348.3

Bioluminescence activity of WT, x2 and x11FLuc. Values within brackets denote the percentage activity retained as compared to 500□M LH_2/ 1mM ATP.

Figure 4.2. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Controls under 500µM LH_2

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	2	1.80114E+15	9.00571E+14	47.35	0.000
Error	6	1.14113E+14	1.90188E+13		
Total	8	1.91526E+15			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
x2	3	47009667	A
x11FLuc	3	29399000	B
WT	3	12359226	C

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from Control Variants under 500µM LH_2/ 1mM ATP. A statistically significant difference between groups was determined by a one-way ANOVA, F(2,6) =47.35, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. N=3

Figure 4.3. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Controls under 20µM LH_2

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	2	1.82033E+12	9.10166E+11	179.81	0.000
Error	5	25309807298	5061961460		
Total	7	1.84564E+12			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
x2 20uM Luc	3	1296500	A
x11FLuc 20uM Luc	3	626920	B
WT 20uM Luc	2	86943	C

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from Control Variants under 20µM LH_2/ 1mM ATP. A statistically significant difference between groups was determined by a one-way ANOVA, F(2,5) =179.81, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. N=3

Figure 4.4. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Controls under 2mM PPi

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	2	1.03614E+13	5.18072E+12	12.39	0.007
Error	6	2.50844E+12	4.18073E+11		
Total	8	1.28699E+13			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
WT 2mM Ppi	3	3734667	A
x2 2mM Ppi	3	3351467	A
x11FLuc 2mM Ppi	3	1291270	B

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from Control Variants under 2mM PPi. A statistically significant difference between groups was determined by a one-way ANOVA, $F(2,6) = 12.39$, $p = 0.007$. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. N=3

4.3.3. Loop Deletion Mutant 'Fingerprinting': A 96-Well Format Screen for Facile Isolation of Novel and Useful Mutants

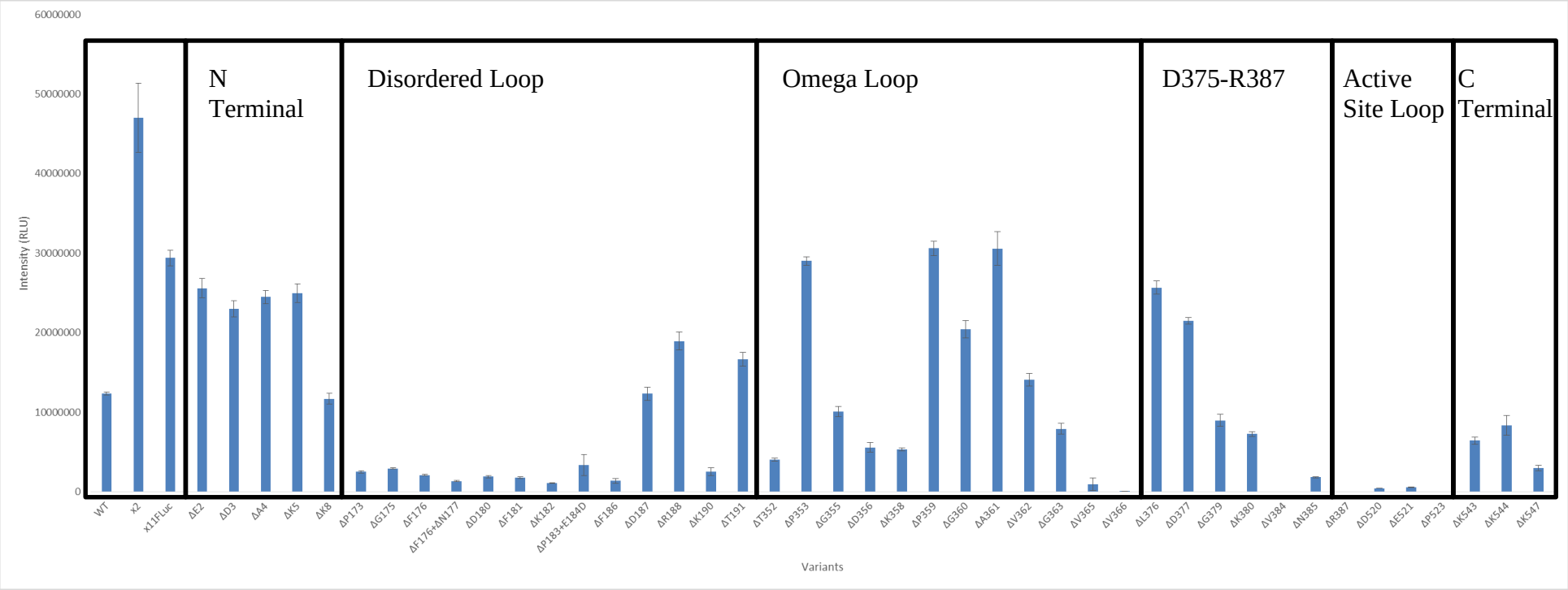
Using the 96-well format, the deletion library of x11Fluc was assayed at RT with saturating LH_2 in triplicate by injecting luminometry with 100ms integration, which only measures the short period of the flash (indicator of Imax). The results of this assay are illustrated in Figure 4.2.

The N-terminal deletions mirrored x11Fluc (29399000RLU), exhibiting activities between 22995000–25600667 RLU (78%-87% of x11FLuc) with the exception of ΔK8 (11708000 RLU, 40% of x11FLuc). The C-terminal deletions exhibited activities of between 2961733-8341067 RLU (11%-28% of x11FLuc). With regards to loops L172-T191 deletions exhibited activities between 1094600–18960667 RLU (3.7%-64% of x11FLuc). T352-F368 deletions exhibited activities between 5316867–30609000 RLU (18%-104% of x11FLuc). D375-R387 deletions exhibited activities between 7753 – 25667667 RLU (0.0002%-87% of x11FLuc). D520-L526 deletions exhibited activities between 2961735–8341067RLU (10%-28% of x11FLuc).

This assay confirmed findings from Chapter 3, confirming the tolerance of x11FLuc to single amino acid deletions, whereby all bar two variants largely retained activity and the activity of these two (□V365 and □V366) was quantified as 976373 RLU (3.3% of x11FLuc) and 93067 RLU (0.3% of x11FLuc), respectively. Highlighting the sensitivity of this method compared to *in vivo* colony screening, and several mutants displayed reduced flash heights.

A clear trend (a 'fingerprint' pattern of activity) was seen throughout screening which broadly correlated to observations in colony screens within Chapter 3 (Figure 4.5.). Although some variation was noted, there was a trend whereby N-terminal deletions were similar to x11FLuc, suggesting that the N-terminal does not play a significant role in modulating activity compared to other regions. However, whilst *in vivo* colony screening determined that C- terminal deletions enhanced activity, the 96-well format concluded that the C-terminal deletions reduced activity. In previous studies making deletions in eGFP (Arpino *et al.,* 2014), both terminals were seen to have similar

Figure 4.5. Bioluminescence Activity of Crude Cell lysates under Saturating Substrate Conditions (Settle Time 0s)



Bioluminescence Activity of Variants under Saturating Substrate Conditions. Crude cell lysates, normalised by OD_{600} of *E.coli BL21* containing WTFLuc, x2FLuc and x11FLuc and deletion variants, induced for 6 hours at RT with IPTG. Bioluminescenet activity measured within Fluoroskan Ascent following injection of 500□M LH_2, 1mM ATP.

Figure 4.6. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between x11FLuc and Single Amino Acid Deletion Mutants

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	43	1.37137E+16	3.18923E+14	233.61	0.000
Error	88	1.20138E+14	1.36520E+12		
Total	131	1.38338E+16			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔP359	3	30609000	A
ΔA361	3	30560667	A
x11FLuc	**3**	**29399000**	**A B**
ΔP353	3	29010000	A B
ΔL376	3	25667667	B C
ΔE2	3	25600667	B C
ΔK5	3	24936333	C D
ΔA4	3	24478000	C D
ΔD3	3	22995000	C D E
ΔD377	3	21466000	D E F
ΔG360	3	20448000	E F G
ΔR188	3	18960667	F G
ΔT191	3	16656333	G H
ΔV362	3	14103000	H I
ΔD187	3	12342000	I J
ΔK8	3	11708000	I J K
ΔG355	3	10099100	J K L
ΔG379	3	8976500	J K L M
ΔK544	3	8341067	K L M
ΔG363	3	7932433	K L M N
ΔK380	3	7265533	L M N
ΔK543	3	6451700	L M N O
ΔD356	3	5593200	M N O P
ΔK358	3	5316867	M N O P
ΔT352	3	4082133	N O P Q
ΔP183+E184D	3	3365600	O P Q R
ΔK547	3	2961733	O P Q R
ΔG175	3	2940833	O P Q R
ΔP173	3	2536267	P Q R
ΔK190	3	2528767	P Q R
ΔF176	3	2121033	P Q R
ΔD180	3	1910333	P Q R
ΔN385	3	1830633	P Q R
ΔF181	3	1815067	P Q R
ΔF186	3	1400227	Q R
ΔF176+ΔN177	3	1345033	Q R
ΔK182	3	1094600	Q R
ΔV365	3	976373	Q R
ΔE521	3	578757	Q R
ΔD520	3	432790	Q R
ΔV366	3	93067	R
ΔP523	3	31640	R
ΔR387	3	7753	R
ΔV384	3	2000	R

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from x11FLuc and deletion mutants under normal conditions. A statistically significant difference between group was determined by a one-way ANOVA, F(43, 88) = 233.61, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc is highlighted in bold. N=3.

tolerance levels to deletion but this does not appear to be the case for FLuc. This may be because the C-terminal is in close proximity to the active site of FLuc which is not the case of eGFP. This observation is exemplified by the reduction in the flash noted for loop D520-L526. Deletions occurring within T352-F368 and D375-R387 displayed the highest Imax values in crude lysates as compared to x11FLuc. Interestingly, ΔP359 and ΔA361 were marginally enhanced with regards to Imax within the initial 100ms.

Furthermore, a double deletion mutant ΔF176+ΔN177 retained similar activity to single amino acid deletions within the same loop. These results suggest that x11FLuc may be able to tolerate larger-scale deletions.

Statistical analysis was conducted to determine the significance of the differences observed (Figure 4.6.). A statistically significant difference between x11FLuc and the deletion variants was found by a one-way ANOVA, $F(43, 88) = 233.61$, $p = 0.000$. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is a significant difference in activity when comparing x11FLuc to the majority of deletion mutants apart from ΔP359, ΔA361, ΔP353, ΔL376 and ΔE2 which display no significant difference.

4.3.4. Fingerprinting the x11Fluc Loop Deletion Library to Identify Mutants with Higher Integrated Activities

To investigate the deletion mutants with regards to integrated activity, following injection of substrate, a reading was measured following a 5 minute settle time with an integration of 100ms using the Fluoroskan instrument. The results of this assay are illustrated in Figure 4.7-4.8. This data indicates the amount of steady-state activity that the enzymes produce after the flash and may help identify mutants with higher specific activities and less decay of signal. This is advantageous in ATP diagnostic assays, which are typically conducted by integration of light over longer periods than the flash.

All mutants displayed a decay in activity over time, however to different degrees. Certain deletion variants displayed higher integrated activity compared to x11FLuc, including ΔK8, ΔP359, ΔV362, ΔG363, ΔV365, ΔV384, R387 and ΔP523 which

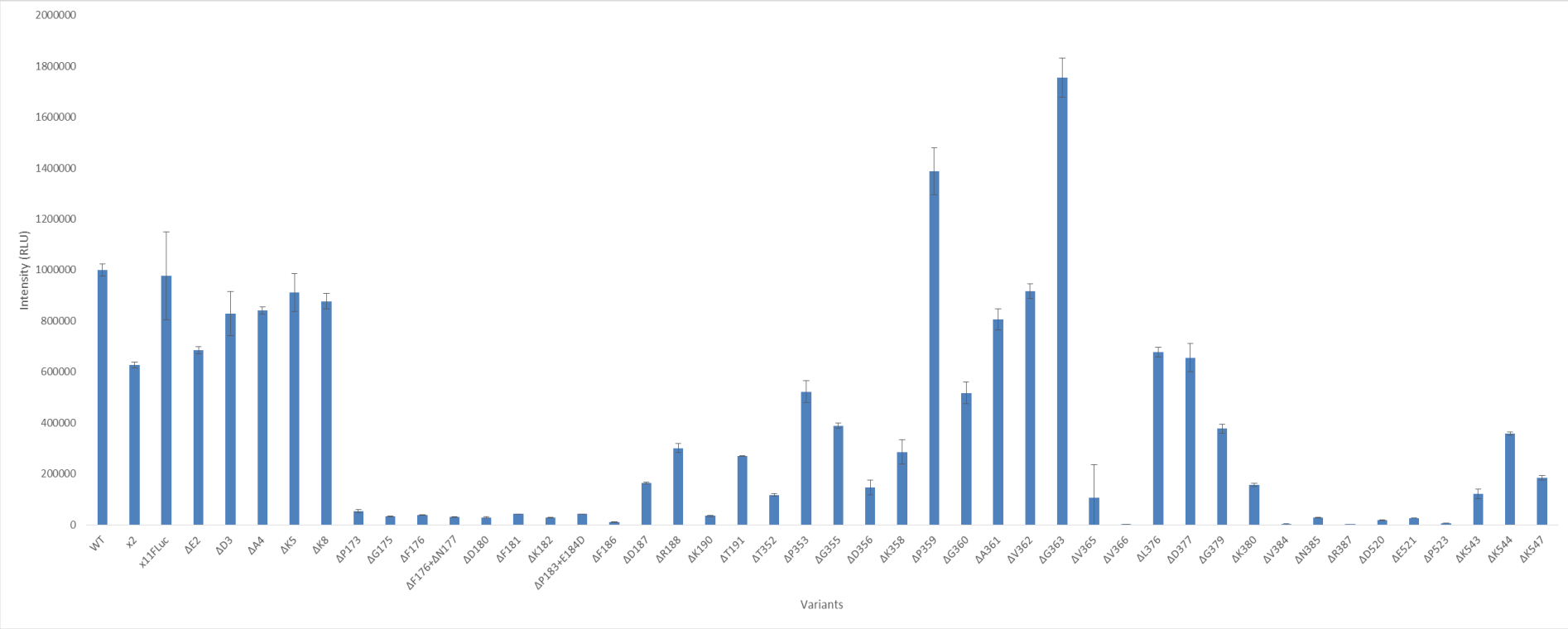
showed 7.49%, 4.53%, 6.49%, 22%, 10%, 76%, 6.4% and 22% decay of initial activities, respectively.

Again, a clear trend was seen throughout screening which correlated better with colony screens than the flash height 96-well assay (section 4.3.5.). This is primarily because *in vivo* colony screening rarely capture the flash which occurs within the first 300ms of the reaction, but captures the overall activity over the 30s of integration.

Eight single amino acid deletion mutants display less decay in activity compared to x11FLuc and include amino acid deletion mutants ΔK8, ΔP359, ΔV362, ΔG363, ΔV365, ΔV384, R387 and ΔP523. This suggests that these mutants either utilise substrates at a slow rate, have less product inhibition or higher specific activity than the parental x11FLuc. It is interesting that 4 of these mutants are derived from the Ω loop.

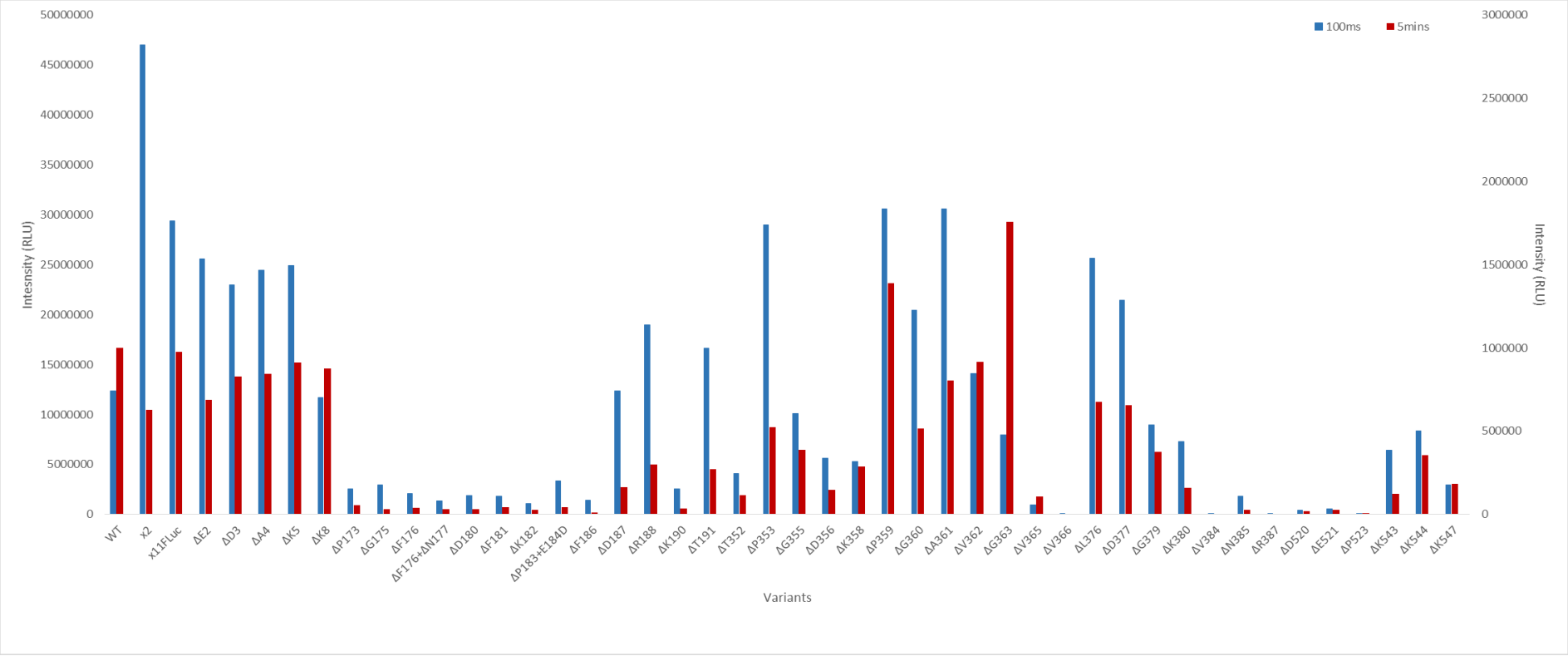
Statistical analysis was conducted to determine the significance of the differences observed x11FLuc and deletion mutants were bioluminescent output was measured following 5000ms (Figure 4.9). A statistically significant difference between x11FLuc and the deletion variants was found by a one-way ANOVA, $F(43, 88) = 119.88$, $p = 0.000$. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is a significant difference in activity when comparing x11FLuc to the majority of deletion mutants apart from ΔV362, ΔK5, ΔK8, ΔA4, ΔD3 and ΔA361 which display no significant difference.

Figure 4.7. Bioluminescence Activity of of Crude Cell lysatess with Saturating Substrate Conditions (Settle Time 5mins)



Bioluminescence Activity of Variants under Saturating Substrate Conditions. Details as in Figure 4.5., but light measured for 100ms after a 5 minute settle time.

Figure 4.8. Bioluminescence Activity of of Crude Cell lysatess with Saturating Substrate Conditions (Settle Time 5mins) compared to Saturating Substrate Conditions (Settle Time 0mins)



Bioluminescence Activity of Variants under Saturating Substrate Conditions. Details as in Figure 4.7.

Figure 4.9. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Activity of x11FLuc and Single Amino Acid Deletion Mutants after a Settle Time of 5000ms

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	43	2.25766E+13	5.25038E+11	119.88	0.000
Error	88	3.85420E+11	4379768772		
Total	131	2.29621E+13			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔG363	3	1755533	A
ΔP359	3	1386867	B
x11FLuc	**3**	**977433**	**C**
ΔV362	3	916557	C
ΔK5	3	911717	C
ΔK8	3	877367	C D
ΔA4	3	842087	C D E
ΔD3	3	829407	C D E
ΔA361	3	805473	C D E
ΔE2	3	685913	D E F
ΔL376	3	676970	D E F
ΔD377	3	656143	E F
ΔP353	3	522850	F G
ΔG360	3	516803	F G H
ΔG355	3	388793	G H I
ΔG379	3	377460	G H I J
ΔK544	3	357320	G H I J K
ΔR188	3	301047	H I J K L
ΔK358	3	286113	I J K L
ΔT191	3	270910	I J K L M
ΔK547	3	184847	I J K L M N
ΔD187	3	163483	J K L M N
ΔK380	3	157407	J K L M N
ΔD356	3	147280	K L M N
ΔK543	3	121409	L M N
ΔT352	3	117313	L M N
ΔV365	3	106415	L M N
ΔP173	3	53896	M N
ΔP183+E184D	3	42352	N
ΔF181	3	42170.3	N
ΔF176	3	39166	N
ΔK190	3	35107	N
ΔG175	3	32948	N
ΔF176+ΔN177	3	30909	N
ΔD180	3	29457	N
ΔN385	3	28463	N
ΔK182	3	28095	N
ΔE521	3	26594	N
ΔD520	3	17507	N
ΔF186	3	10529	N
ΔP523	3	6911	N
ΔV384	3	1520	N
ΔV366	3	1516.0	N
ΔR387	3	499.3	N

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from x11FLuc and deletion mutants under normal conditions measured after 5000ms. A statistically significant difference between group was determined by a one-way ANOVA, F(43, 88) = 119.88, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc is highlighted in bold. N=3.

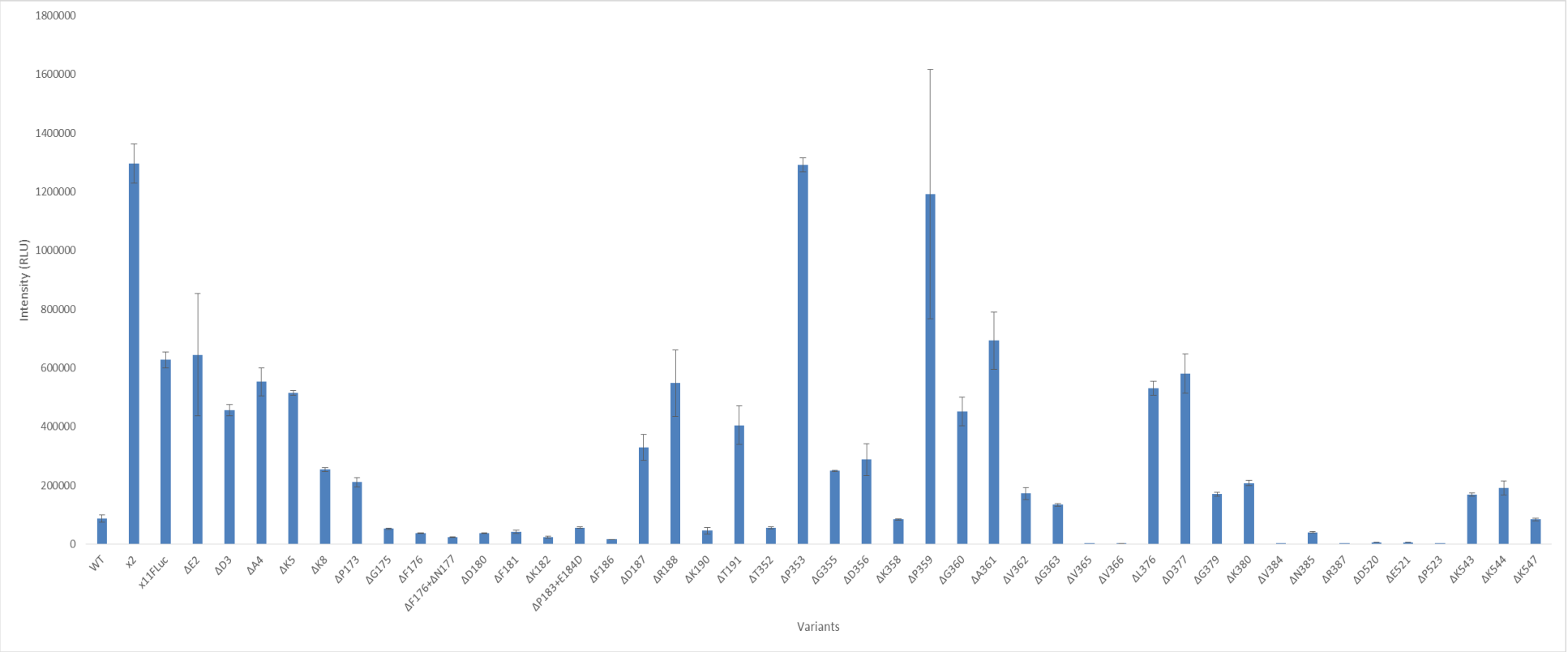
4.3.5. Fingerprinting the x11Fluc Loop Deletion Library to Identify Mutants for Lower K_M for D-LH_2

To screen the deletion mutants for lower K_M for LH_2, 20□M LH_2 (non-saturating concentration of LH_2) substrate was injected and light measured with an integration of 100ms and results were compared to the results obtained in section 4.3.3 (saturating concentration of LH_2). The results of this assay are illustrated in Figure 4.10-4.11.

Four mutants in particular displayed the potential for a lower K_M for LH_2, those being ΔP173, ΔP353, ΔD356 and ΔP359. In comparison to x11FLuc whose activity was 2.13% under substrate limiting conditions, ΔP173, ΔP353, ΔD356 and ΔP359 displayed activities of 8.3%, 4.4%, 5.14%, 3.89%, respectively.

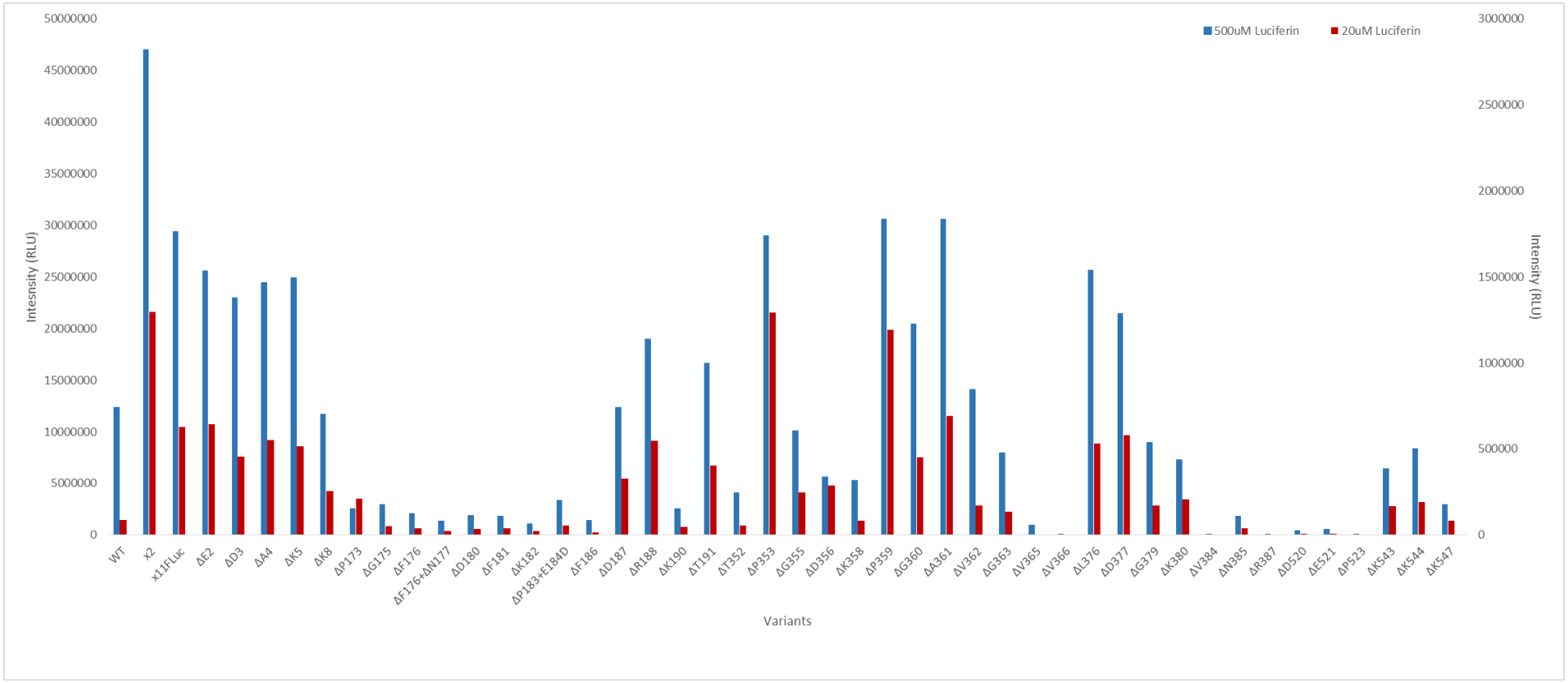
It appears that of the single amino acid deletion mutants, particularly, 4 appeared to have the potential to achieve higher velocity at low concentrations of the LH_2 substrate in comparison to the controls, these include ΔP173, ΔP353, ΔD356 and ΔP359. It is of note that 3 of these potential candidates were deletions derived from the Ω loop. Statistical analysis was conducted to determine the significance of the differences observed x11FLuc and deletion mutants when subject to limiting conditions of the substrate (Figure 4.12). A statistically significant difference between x11FLuc and the deletion variants was found by a one-way ANOVA, $F(43, 88) = 22.55$ $p = 0.000$. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is a significant difference in activity when comparing x11FLuc to the majority of deletion mutants apart from ΔA361, ΔE2, ΔD377, ΔA4, ΔR188, ΔL376, ΔK5, ΔD3, ΔG360, ΔT191, D187 which display no significant difference.

Figure 4.10. Bioluminescence Activity of Variants under Conditions of Non Saturating LH_2



Bioluminescence Activity of Variants under Non Saturating Substrate Conditions. Details as in Figure 4.5., but light measured for 100ms after injeciton of 20□M LH_2 and 1mM ATP.

Figure 4.11. Bioluminescence Activity of Variants under Conditions of Non Saturating LH_2 compared to Saturating Substrate Conditions (Settle Time 0mins)



Bioluminescence Activity of Variants under Non-Saturating Substrate Conditions. Details as in Figure 4.10.

Figure 4.12. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Activity of x11FLuc and Single Amino Acid Deletion Mutants under Non-Saturating Conditions

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	43	1.20968E+13	2.81321E+11	22.55	0.000
Error	88	1.09775E+12	12474466428		
Total	131	1.31946E+13			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔP353	3	1291533	A
ΔP359	3	1191813	A
ΔA361	3	693337	B
ΔE2	3	644780	B C
x11FLuc	**3**	**626920**	**B C**
ΔD377	3	580097	B C D
ΔA4	3	552160	B C D E
ΔR188	3	547723	B C D E
ΔL376	3	530253	B C D E F
ΔK5	3	514737	B C D E F
ΔD3	3	456070	B C D E F G
ΔG360	3	452253	B C D E F G H
ΔT191	3	404307	B C D E F G H I
ΔD187	3	329010	B C D E F G H I J
ΔD356	3	288003	C D E F G H I J
ΔK8	3	253653	D E F G H I J
ΔG355	3	248977	D E F G H I J
ΔP173	3	210767	D E F G H I J
ΔK380	3	207343	E F G H I J
ΔK544	3	191250	E F G H I J
ΔV362	3	171687	F G H I J
ΔG379	3	169727	F G H I J
ΔK543	3	168120	F G H I J
ΔG363	3	133530	G H I J
ΔK358	3	84263	G H I J
ΔK547	3	83447	H I J
ΔP183+E184D	3	56066	I J
ΔT352	3	55339	I J
ΔG175	3	51958	I J
ΔK190	3	45273	I J
ΔF181	3	41207	I J
ΔN385	3	39892	I J
ΔF176	3	37324	I J
ΔD180	3	36307	I J
ΔF176+ΔN177	3	23049	J
ΔK182	3	22946	J
ΔF186	3	15537	J
ΔD520	3	5839	J
ΔE521	3	5743	J
ΔV366	3	2497.0	J
ΔP523	3	309.3	J
ΔV365	3	40.0	J
ΔV384	3	39.33	J
ΔR387	3	16.0	J

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from x11FLuc and deletion mutants under non-saturating conditions. A statistically significant difference between group was determined by a one-way ANOVA, $F(43, 88) = 22.55$, $p = 0.000$. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc is highlighted in bold. N=3.

4.3.6. Fingerprinting the x11FLuc Loop Deletion Library to Identify Mutants for Resistance to Inhibition by Inorganic Pyrophosphate

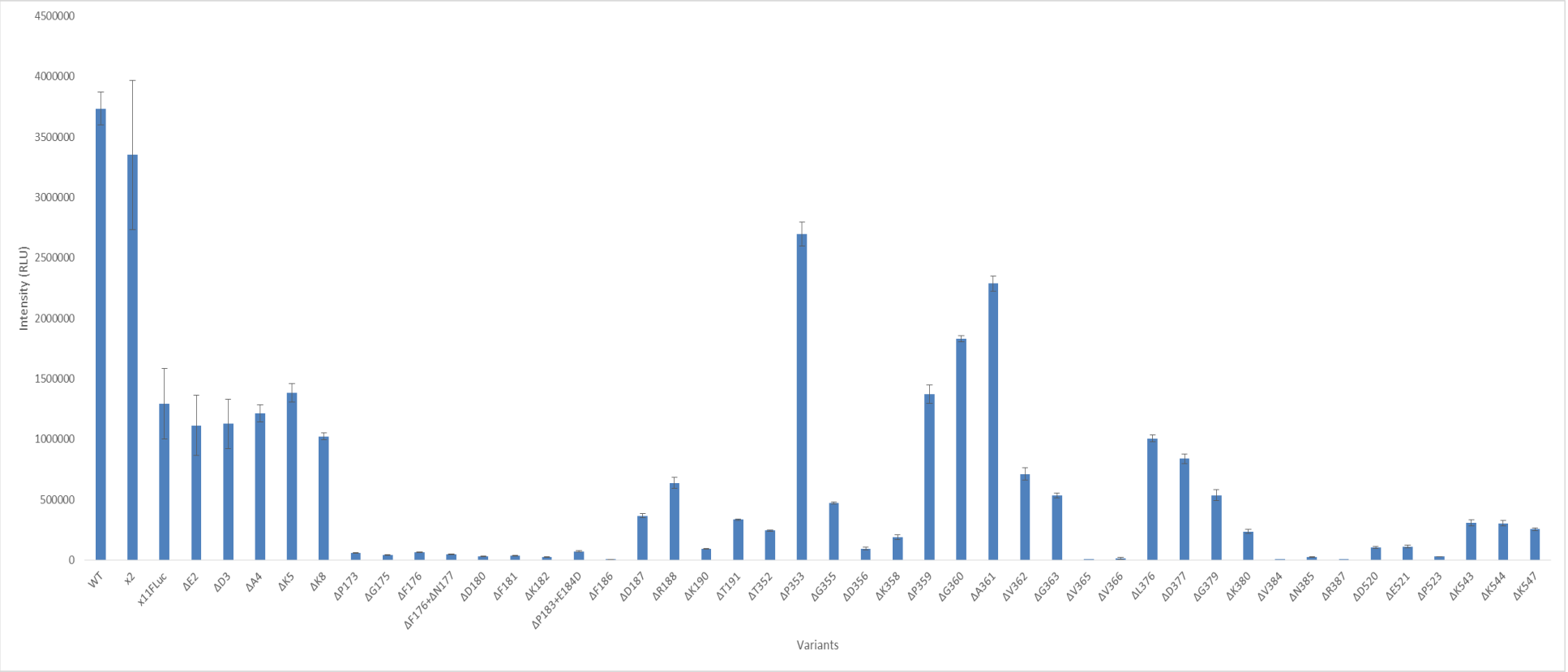
To investigate the deletion mutants with regards to resistance to inhibition by 2mM inorganic pyrophosphate (PPi), following injection of LH_2 substrate light emission was measured with an integration of 100ms and results compared to those obtained in section 4.3.3. The results of this assay are illustrated in Figure 4.13-4.14

Three mutants (ΔV384, ΔD520 and ΔP523) displayed a particularly high resistance to PPi. Their activities were 36%, 23% and 86% in the presence of PPi compared to without it, compared to 4.4% for x11FLuc. In addition to this, certain of the Ω loop variants also displayed an enhanced resistance to PPi compared to x11FLuc, there being ΔP353, ΔG360, ΔA361 and ΔG363, whose activities are 9.2, 8.95, 7.4 and 6.72%, respectively, with PPi compared to without it. Utilising this strategy, many deletion mutants have been identified which appear to be more resistant to PPi inhibition than x11FLuc.

The crude cell lysate strategy correlates well with the levels of inhibition of FLuc by PPi reported in the literature (Fontes *et al*., 2008). The presence of the PPi inhibits bioluminescence due to the reversible nature of the adenylation reaction. Le Châtliers principle states that high concentrations of PPi will reverse the direction of the reaction toward the left (see Chapter 1, Equation 1.1.).

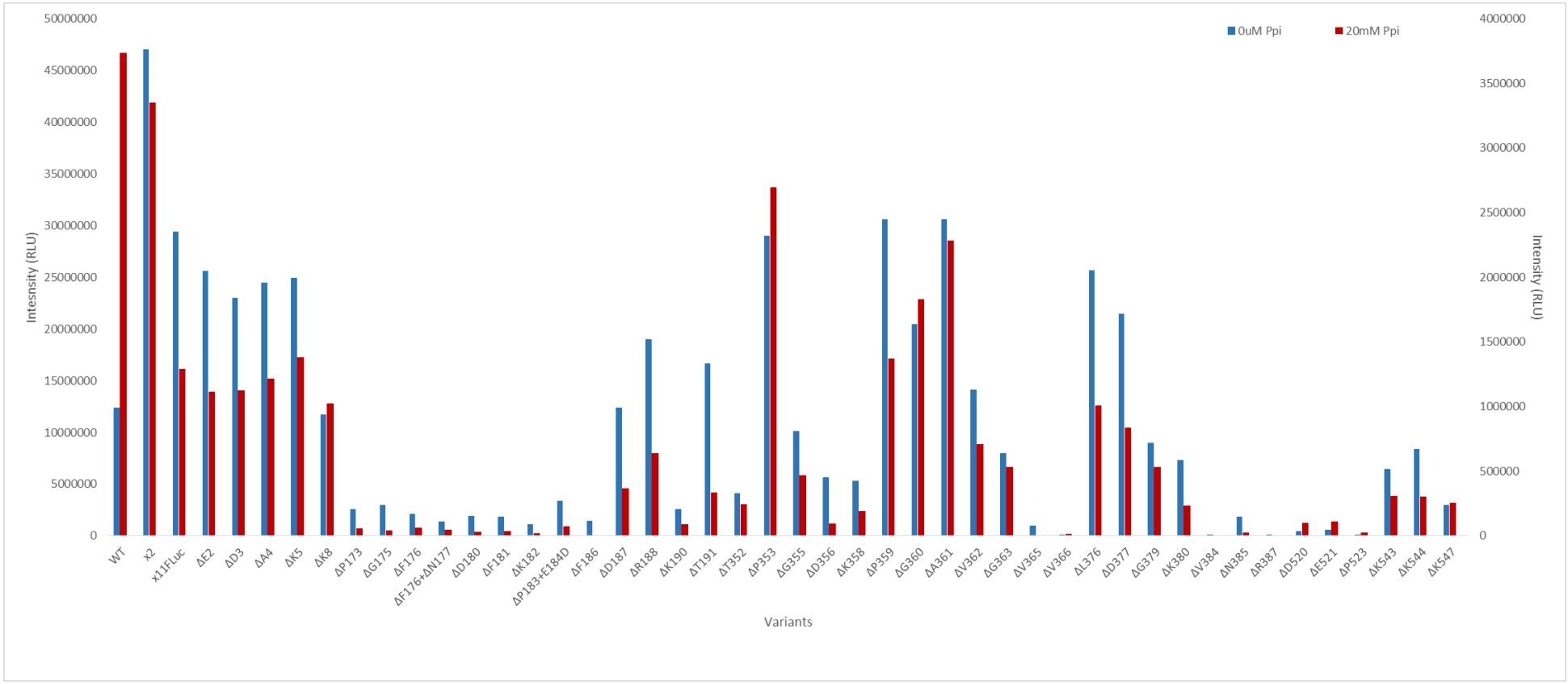
Statistical analysis was conducted to determine the significance of the differences observed x11FLuc and deletion mutants when subject to PPi (Figure 4.15). A statistically significant difference between x11FLuc and the deletion variants was found by a one-way ANOVA, $F(43, 88) = 95.59$, $p = 0.000$. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is a significant difference in activity when comparing x11FLuc to the majority of deletion mutants apart from ΔK5, ΔP359, ΔA4, ΔD3, ΔE2, ΔK8, ΔL376 which display no significant difference.

Figure 4.13. Bioluminescence Activity of Variants under Conditions of PPi



BioluminesceneceActivity of Variants under Saturating Substrate Conditions of PPi. Details as in Figure 4.5., but light measured for 100ms after addition of 2mM PPi.

Figure 4.14. Bioluminescence Activity of Variants under Conditions of PPi compared to Saturating Substrate Conditions (Settle Time 0mins)



BioluminesceneceActivity of Variants under Saturating Substrate Conditions of PPi. Details as in Figure 4.13.

Figure 4.15. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Activity of x11FLuc and Single Amino Acid Deletion Mutants under PPi

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	43	5.48991E+13	1.27672E+12	95.59	0.000
Error	88	1.17535E+12	13356247151		
Total	131	5.60745E+13			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔP353	3	2696900	A
ΔA361	3	2286567	B
ΔG360	3	1830500	C
ΔK5	3	1383300	D
ΔP359	3	1372100	D
x11FLuc	**3**	**1291270**	**D**
ΔA4	3	1215067	D E
ΔD3	3	1125770	D E
ΔE2	3	1113533	D E
ΔK8	3	1022777	D E F
ΔL376	3	1005420	D E F
ΔD377	3	837390	E F G
ΔV362	3	711293	F G H
ΔR188	3	637793	F G H I
ΔG379	3	535057	G H I J
ΔG363	3	533577	G H I J
ΔG355	3	469810	G H I J K
ΔD187	3	365573	H I J K L
ΔT191	3	333707	H I J K L
ΔK543	3	307530	I J K L
ΔK544	3	304397	I J K L
ΔK547	3	254703	I J K L
ΔT352	3	243893	J K L
ΔK380	3	236253	J K L
ΔK358	3	188537	J K L
ΔE521	3	109936	K L
ΔD520	3	102860	K L
ΔD356	3	93718	K L
ΔK190	3	91737	K L
ΔP183+E184D	3	71544	L
ΔF176	3	64783	L
ΔP173	3	57622	L
ΔF176+ΔN177	3	47994	L
ΔG175	3	40233	L
ΔF181	3	34683	L
ΔD180	3	31470	L
ΔP523	3	27291	L
ΔN385	3	23113	L
ΔK182	3	22835	L
ΔV366	3	13038	L
ΔF186	3	4698	L
ΔV384	3	736	L
ΔV365	3	273	L
ΔR387	3	192.7	L

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from x11FLuc and deletion mutants under non-saturating conditions. A statistically significant difference between group was determined by a one-way ANOVA, $F(43, 88) = 95.59$, $p = 0.000$. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc is highlighted in bold. N=3.

4.3.7. Fingerprinting the x11FLuc Loop Deletion Library to Identify Mutants for Resistance to Thermal Inactivation

To investigate the deletion mutants with regards to those potentially exhibiting resistance to thermal inactivation, light emission was measured for an integration time of 100ms post LH_2 injection following incubation at RT, 42ºC and 60ºC. The results of this assay are illustrated in Figure 4.6.

WTFluc displayed a large decrease in activity at 42ºC to 60ºC, with activities of 19% to 0.006% compared to at RT (100%), respectively, whilst x11FLuc exhibited activities of 237 to 142%, which was increased compared to the RT value. The N-terminal deletions mirrored x11FLuc, exhibiting activities between 204-252% at 42ºC and between 123-146% at 60ºC, with the exception of ΔK8 which reduced to 40% activity at 60ºC. The C-terminal deletions exhibited activities between 277-352% at 42ºC and between 140-163% at 60ºC. With regards to loops, L172- T191 deletions exhibited activities between 163%-301% at 42ºC, with the exception of ΔK182 and ΔF186 (95% and 18%, respectively) and between 107%-242% at 60ºC with the exception of □P183/ E184D, □F186 and □D187 (2.09, 0.09 and 21% respectively). T352-F368 loop deletions exhibited activities between 107%-252% at 42ºC, with the exception of ΔK358 and between 0.08%-76% at 60ºC with the exception of ΔT353 (118%). D375-R387 loop deletions exhibited activities between 166%-327% at 42ºC, with the exception of ΔR387 (23%) and between 147%-325% at 60ºC with the exception of ΔK380, ΔN385 and ΔR387 (99%, 5% and 0.82% respectively). D520-L526 deletions exhibited activities between 148%-155% at 42ºC with the exception of ΔP523 (52%) and between 31%-89% at 60ºC.

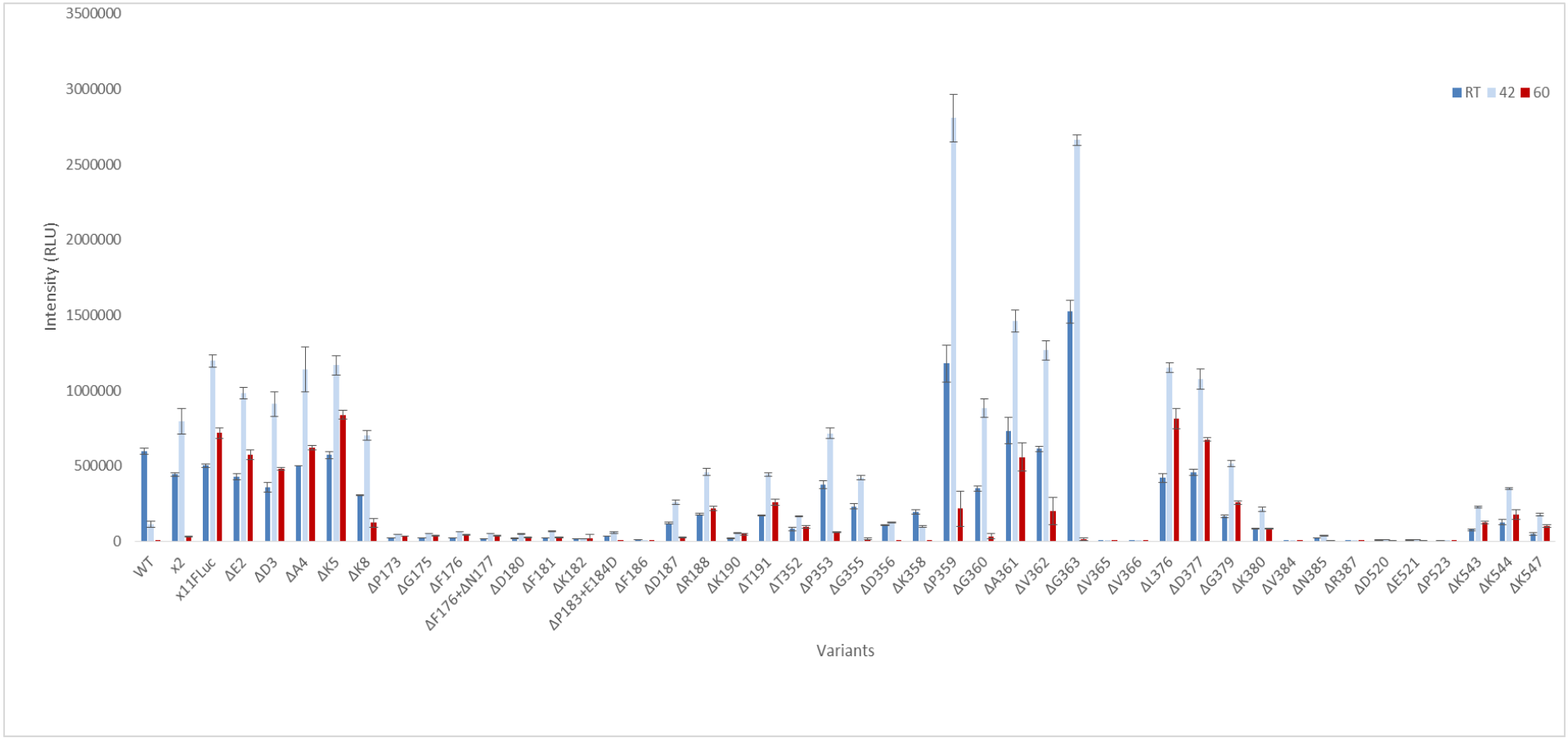
In this assay, the basic fingerprint of loop activities was the similar to the results of ImageJ of secondary screens (see Chapter 3, section 3.3.8.3.). Whilst deletions within the N- and C-termini were similar in terms of resistance to thermal inactivation to x11FLuc, remarkably, deletions within loop L172- T191 appeared to confer greater stability than that of x11FLuc. Of all the loops investigated, deletion mutants in the Ω loop T352-F368 displayed the greatest losses in activity at 60ºC, although their stability still appeared higher than that of either the WTFLuc or x2FLuc. Therefore, although mutants have altered peptide backbones, the 11 mutations involved in the stabilisation

of x11FLuc still confer some stabilisation. On the other hand, deletions in loop D375-R387, which is in close proximity downstream of the Ω loop, result in variations in resistance to thermal inactivation. Deletions promoted enhanced activity at 60ºC as when located towards the N-terminal of this loop, whilst when towards the C-terminal of the loop, they exhibited the opposite characteristic, displaying less activity at 60ºC compared to RT. Lastly, the resistance of deletion mutants within the active site loop D520-L526 were similar to the Ω loop and were less stable at 60ºC compared to x11FLuc. Indeed, ΔP523 retained less activity following incubation at 42ºC suggesting that deletions close to active site destabilised the protein.

It is of interest that whilst overall activity did not correlate well with conservation scoring (as determined in Chapter 3), conservation scoring does appear to correlate with the toleration of the deletions with regards to resistance to thermal inactivation. The loops presenting high conservation scores T352-F368 and D520-L526 (Chapter 3, Table 3.3.) displayed the least resistance to thermal inactivation, whereas deletions within those loops with a low conservation score L172- T191 appear to promote further stabilisation of the enzyme.

Statistical analysis was conducted to determine the significance of the differences observed x11FLuc and deletion mutants at 42ºC and 60ºC (Figure 4.17-4.18). A statistically significant difference between x11FLuc and the deletion variants was found by a one-way ANOVA at both temperatures. The Tukey method was further employed to group variants by means that are not significantly different to confirm observations such that there is a significant difference in activity when comparing x11FLuc to the majority of deletion mutants at 42ºC apart from ΔV362, ΔK5, ΔL376, ΔA4, ΔD377, ΔE2, ΔD3 which display no significant difference. At 62ºC, there is a significant difference in activity when comparing x11FLuc to the majority of deletion mutants apart from apart from apart from ΔE5, ΔL376, ΔD377, ΔA4, ΔE2, ΔA362 which display no significant difference.

Figure 4.16. Bioluminescence Activity of Variants Incubated at 40 ºC and 60ºC



Bioluminescence Activity of Variants following Incubation at RT, 42ºC and 60ºC. Details as in Figure 4.5., but light measured following incubation at RT, 42 ºC and 60ºC.

Figure 4.17. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Activity of x11FLuc and Single Amino Acid Deletion Mutants at 42 ºC

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	43	5.69996E+13	1.32557E+12	293.02	0.000
Error	88	3.98103E+11	4523897218		
Total	131	5.73977E+13			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔP359	3	2807933	A
ΔG363	3	2660200	A
ΔA361	3	1460433	B
ΔV362	3	1267833	B C
x11FLuc	**3**	**1196633**	**C D**
ΔK5	3	1167967	C D
ΔL376	3	1151667	C D
ΔA4	3	1139133	C D
ΔD377	3	1077980	C D E
ΔE2	3	980963	D E
ΔD3	3	911350	E F
ΔG360	3	885170	E F
ΔP353	3	714880	F G
ΔK8	3	703893	F G
ΔG379	3	517937	G H
ΔR188	3	459660	H I
ΔT191	3	444263	H I J
ΔG355	3	422753	H I J K
ΔK544	3	351587	H I J K L
ΔD187	3	260337	I J K L M
ΔK543	3	228567	J K L M N
ΔK380	3	212950	K L M N O
ΔK547	3	177157	L M N O
ΔT352	3	164450	L M N O
ΔD356	3	126997	M N O
ΔK358	3	97217	M N O
ΔF181	3	67653	M N O
ΔF176	3	63742	M N O
ΔP183+E184D	3	58215	M N O
ΔK190	3	56109	M N O
ΔF176+ΔN177	3	53185	M N O
ΔG175	3	51409	M N O
ΔD180	3	49159	M N O
ΔP173	3	46571	M N O
ΔN385	3	36838	M N O
ΔK182	3	17630	N O
ΔE521	3	12881	N O
ΔD520	3	12079	N O
ΔP523	3	2240.0	O
ΔV366	3	2093	O
ΔF186	3	1957	O
ΔV365	3	140.0	O
ΔV384	3	120.0	O
ΔR387	3	96.7	O

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from x11FLuc and deletion mutants at 42ºC. A statistically significant difference between group was determined by a one-way ANOVA, F(43, 88) =293.02, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc is highlighted in bold. N=3.

Figure 4.18. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between Activity of x11FLuc and Single Amino Acid Deletion Mutants at 60 ºC

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	43	7.99218E+12	1.85865E+11	76.00	0.000
Error	88	2.15204E+11	2445496904		
Total	131	8.20739E+12			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔK5	3	839767	A
ΔL376	3	814357	A
x11FLuc	**3**	**718370**	**A B**
ΔD377	3	675173	A B
ΔA4	3	621323	B C
ΔE2	3	575060	B C
ΔA361	3	558030	B C
ΔD3	3	479643	C
ΔT191	3	260107	D
ΔG379	3	256847	D
ΔR188	3	218337	D E
ΔP359	3	215686	D E
ΔV362	3	203176	D E F
ΔK544	3	178260	D E F G
ΔK543	3	126220	D E F G H
ΔK8	3	123831	D E F G H
ΔK547	3	101676	D E F G H
ΔT352	3	96641	D E F G H
ΔK380	3	83966	E F G H
ΔP353	3	59760	E F G H
ΔK190	3	47347	F G H
ΔF176	3	44022	F G H
ΔG175	3	37653	G H
ΔF176+ΔN177	3	35738	G H
ΔP173	3	35667	G H
ΔG360	3	34127	G H
ΔF181	3	26896	G H
ΔD187	3	25667	G H
ΔD180	3	25601	G H
ΔK182	3	19803	G H
ΔG355	3	16797	G H
ΔG363	3	13937	G H
ΔE521	3	7421	H
ΔD520	3	6550	H
ΔN385	3	3467	H
ΔP523	3	1352.7	H
ΔP183+E184D	3	740	H
ΔV366	3	242.7	H
ΔD356	3	203.3	H
ΔK358	3	166.7	H
ΔV384	3	119.3	H
ΔV365	3	20.0	H
ΔF186	3	10.00	H
ΔR387	3	3.33	H

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from x11FLuc and deletion mutants at 60ºC. A statistically significant difference between group was determined by a one-way ANOVA, F(43, 88) = 78, p = 0.000. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc is highlighted in bold. N=3.

Table 4.2 Summary of Maximal Activity of Variants in Different Assay Conditions as a Percentage

	Bioluminescence Activity (RLU)					
	0-100ms (RLU)	5 minutes (%)	Low K_M (20µM)(%)	PPi (2mM) (%)	Incubated at 42ºC (%)	Incubated at 60ºC (%)
WT	12359226.00	8.10	0.70	30.22	19.03	0.01
x2	47009667.00	1.33	2.76	7.13	178.70	6.81
x11FLuc	29399000.00	3.32	2.13	4.39	237.84	142.78
ΔE2	25600667.00	2.68	2.52	4.35	229.61	134.60
ΔD3	22995000.00	3.61	1.98	4.90	252.58	132.93
ΔA4	24478000.00	3.44	2.26	4.96	226.66	123.63
ΔK5	24936333.00	3.66	2.06	5.55	204.22	146.83
ΔK8	11708000.00	7.49	2.17	8.74	231.64	40.75
ΔP173	2536267.00	2.13	8.31	2.27	214.71	164.44
ΔG175	2940833.00	1.12	1.77	1.37	226.43	165.84
ΔF176	2121033.00	1.85	1.76	3.05	281.45	194.38
ΔF176+	1345033.00	2.30	1.71	3.57	301.67	202.71
ΔD180	1910333.00	1.54	1.90	1.65	249.15	129.75
ΔF181	1815067.00	2.32	2.27	1.91	301.89	120.02
ΔK182	1094600.00	2.57	2.10	2.09	95.65	107.44
ΔP183+	3365600.00	1.26	1.67	2.13	164.86	2.10
ΔF186	1400227.00	0.75	1.11	0.34	18.57	0.09
ΔD187	12342000.00	1.32	2.67	2.96	213.50	21.05
ΔR188	18960667.00	1.59	2.89	3.36	257.17	122.16
ΔK190	2528767.00	1.39	1.79	3.63	287.01	242.19
ΔT191	16656333.00	1.63	2.43	2.00	258.34	151.25
ΔT352	4082133.00	2.87	1.36	5.97	201.32	118.31
ΔP353	29010000.00	1.80	4.45	9.30	190.75	15.95
ΔG355	10099100.00	3.85	2.47	4.65	182.50	7.25
ΔD356	5593200.00	2.63	5.15	1.68	119.91	0.19
ΔK358	5316867.00	5.38	1.58	3.55	49.98	0.09
ΔP359	30609000.00	4.53	3.89	4.48	238.26	18.30
ΔG360	20448000.00	2.53	2.21	8.95	252.45	9.73
ΔA361	30560667.00	2.64	2.27	7.48	199.12	76.08
ΔV362	14103000.00	6.50	1.22	5.04	207.02	33.18
ΔG363	7932433.00	22.13	1.68	6.73	174.66	0.92
ΔV365	976373.30	10.90	0.00	0.03	107.69	15.38
ΔV366	93067.33	1.63	2.68	14.01	143.38	16.62
ΔL376	25667667.00	2.64	2.07	3.92	273.05	193.07
ΔD377	21466000.00	3.06	2.70	3.90	234.85	147.09
ΔG379	8976500.00	4.20	1.89	5.96	315.83	156.62
ΔK380	7265533.00	2.17	2.85	3.25	252.13	99.41

ΔV384	1999.67	76.01	1.97	36.81	327.27	325.45
ΔN385	1830633.00	1.55	2.18	1.26	166.73	15.69
ΔR387	7753.00	6.44	0.21	2.49	23.97	0.83
ΔD520	432790.00	4.05	1.35	23.77	148.60	80.58
ΔE521	578756.70	4.60	0.99	19.00	155.96	89.85
ΔP523	31640.00	21.84	0.98	86.25	52.94	31.97
ΔK543	6451700.00	1.88	2.61	4.77	295.77	163.33
ΔK544	8341067.00	4.28	2.29	3.65	277.07	140.48
ΔK547	2961733.00	6.24	2.82	8.60	352.32	202.21

Table illustrating the relative levels of activity at different stages of the reaction and in different conditions. Results displayed as a percentage of flash-based activity.

Table 4.3. Summary of Maximal Activity of Variants in Different Assay Conditions as a Percentage of x11FLuc

	Bioluminescence Activity (RLU)					
	0-100ms (RLU)	5 minutes (%)	Low K_M (20µM)(%)	PPi (2mM) (%)	Incubated at 42ºC (%)	Incubated at 60ºC (%)
ΔE2	87.08	70.17	102.85	86.24	81.97694	80.05067
ΔD3	78.22	84.86	72.75	87.18	76.1595	66.76829
ΔA4	83.26	86.15	88.08	94.10	95.19485	86.49071
ΔK5	84.82	93.28	82.11	107.13	97.60439	116.8989
ΔK8	39.82	89.76	40.46	79.21	58.82281	17.23777
ΔP173	8.63	5.51	33.62	4.46	3.891863	4.96499
ΔG175	10.00	3.37	8.29	3.12	4.296109	5.241449
ΔF176	7.21	4.01	5.95	5.02	5.32675	6.128086
ΔF176+	4.58	3.16	3.68	3.72	4.444525	4.974827
ΔD180	6.50	3.01	5.79	2.44	4.108109	3.563716
ΔF181	6.17	4.31	6.57	2.69	5.653612	3.744032
ΔK182	3.72	2.87	3.66	1.77	1.4733	2.756704
ΔP183+	11.45	4.33	8.94	5.54	4.864871	0.103011
ΔF186	4.76	1.08	2.48	0.36	0.163514	0.001392
ΔD187	41.98	16.73	52.48	28.31	21.75576	3.57295
ΔR188	64.49	30.80	87.37	49.39	38.41277	30.39334
ΔK190	8.60	3.59	7.22	7.10	4.688933	6.590893
ΔT191	56.66	27.72	64.49	25.84	37.1261	36.2079
ΔT352	13.89	12.00	8.83	18.89	13.74272	13.45282
ΔP353	98.68	53.49	206.01	208.86	59.74094	8.318787
ΔG355	34.35	39.78	39.71	36.38	35.32856	2.33821
ΔD356	19.03	15.07	45.94	7.26	10.61283	0.028305
ΔK358	18.09	29.27	13.44	14.60	8.124182	0.023201
ΔP359	104.12	141.89	190.11	106.26	234.6528	30.02441
ΔG360	69.55	52.87	72.14	141.76	73.9717	4.750616
ΔA361	103.95	82.41	110.59	177.08	122.0452	77.68003
ΔV362	47.97	93.77	27.39	55.08	105.95	28.28296
ΔG363	26.98	179.61	21.30	41.32	222.307	1.940133
ΔV365	3.32	10.89	0.01	0.02	0.011699	0.002784
ΔV366	0.32	0.16	0.40	1.01	0.174935	0.03378
ΔL376	87.31	69.26	84.58	77.86	96.24224	113.3617
ΔD377	73.02	67.13	92.53	64.85	90.0844	93.98685
ΔG379	30.53	38.62	27.07	41.44	43.28282	35.75409
ΔK380	24.71	16.10	33.07	18.30	17.79576	11.68841
ΔV384	0.01	0.16	0.01	0.06	0.010028	0.016612
ΔN385	6.23	2.91	6.36	1.79	3.078498	0.482667
ΔR387	0.03	0.05	0.00	0.01	0.008078	0.000464

ΔD520	1.47	1.79	0.93	7.97	1.009415	0.911786
ΔE521	1.97	2.72	0.92	8.51	1.076409	1.033033
ΔP523	0.11	0.71	0.05	2.11	0.187192	0.188297
ΔK543	21.95	12.42	26.82	23.82	19.10081	17.57033
ΔK544	28.37	36.56	30.51	23.57	29.38132	24.81451
ΔK547	10.07	18.91	13.31	19.73	14.80459	14.15371

Table illustrating the relative levels of activity at different stages of the reaction and in different conditions. Results displayed as a percentage of x11FLuc activity.

4.4. Further Discussion

The aim of this investigation was to utilise an alternative strategy in order to refine the pool of deletion mutants generated in Chapter 3 to take forward to characterisation. An alternative method based on crude cell lysate screening was developed in an effort to study bioluminescent properties observed in previous chapters. This alternative strategy incorporated a 96-well format to allow for the characterisation of the library of deletion mutants with regards to their kinetics, specific activities, resistance to inhibition and thermal inactivation.

In each assay condition there was a clear and generic trend, which could be broadly correlated to *in vivo* colony screening (Chapter 3), that being, overall activity as a result N-terminal deletions was similar to x11FLuc, whilst C-terminal deletions and deletions within L172- T191 and D520-L526 promoted reductions in activity. On the other hand, deletions within T352-F368 and D375-R387 exhibited wide ranging impacts on activity. In particular, T352-F368 displayed enhancements in comparison to x11FLuc. It is of interest that deletions occurring within disparate loops broadly confer similar characteristics to x11FLuc. However, neighbouring mutations within a single loop, in particular T352-F368, can confer markedly different characteristics.

This assay format allowed assessment of the impact of deletions on various characteristics, such as specific activity, resistance to PPi and may be useful to determine K_M. ΔK8, ΔP359, ΔV362, ΔG363, ΔV365, ΔV384, R387 and ΔP523 were identified as potential mutants with higher specific activity whilst ΔP173, ΔP353, ΔD356 and ΔP359 exhibited potential to have a reduced K_M for LH_2. These results taken together with overall activity highlighted the deletions within the Ω loop generated mutants that were not only enhanced in activity compared to x11Fluc but also displayed colour shifts and lower K_M for LH_2. In addition to this, within Chapter 3, deletions within the Ω loop caused variation in □max and FWHM of bioluminescence spectra from crude lysates. Therefore, due to the high variety of potentially useful phenotypes derived from deletions in the omega loop, and its structural significance, these mutants were taken forward and characterised in Chapter 5.

As with *in vivo* colony screening, there are limitations in the 96-well format screening strategy, which include that differences in the activity between mutants may be due to changes in their expression levels. Conventionally, substitution mutations tend to impart little change in the overall expression of mutants, however, in the literature, deletion mutants have been identified that improve the folding landscape of GFP (Arpino *et al*., 2014). GFP conventionally has a long maturation time, however, a deletion of G4 increased the rate of folding of a protein. In contrast, WTFLuc folds rapidly after translation by ribosomes (Svetlov *et al.*, 2006), therefore, this changes in maturation time are likely to be less important in the case of mutants generated here. A likely scenario, is that there is an increase in activity due to an increase in expression and this may be attributed to better folding and thus increased levels of soluble active protein.

With this assay however, it was not possible to address this issue. *E. coli* containing mutants were normalised by optical density measurement of cells for resulting in good correlation between assays. However, this does not overcome the issue of potentially varied expression fully. Assays could be redesigned to quantify protein expression during screening by linking of a fluorescent protein, such as Cyan Fluorescent Protein (CFP) (Kremer *et al*., 1995) or Enhanced Blue Fluorescent Protein 2 (EBFP2) (Ai *et al.*, 2007) onto the N-terminal of FLuc. Quantification of fluorescence could be utilised to determine the relative expression of FLuc mutants, whilst since the excitation of these fluorescent proteins is blue shifted compared to the emission of FLuc, bioluminescence resonance energy transfer (BRET) will not occur

Another limitation of the strategy is that whilst this strategy is the inability to control levels of ATP. As part of this study, it would be desirable to screen mutants for reductions in ATP K_M under low ATP conditions. However, as crude cell lysates naturally contain ATP these assays must be performed on purified protein.

Despite the limitations of the technique, utilising this strategy has isolated brighter mutants, but whether this is due to a true increase in activity or merely as a result of an increase in expression is not known. Since deletions can promote changes in expression levels, and increased expression levels are desirable, they have the potential to improve mutants and ease of their purification in a way that substitutions cannot. This is particularly important in a manufacturing setting whereby it is desirable to have mutants

which are more likely to fold correctly to maximise the volume of functional enzyme produced.

An advantage of this method is that unlike mutational studies which fully characterise few mutants generated from a library, it was possible to sample many single amino acid deletions within a number of loops simultaneously. This allowed the investigation of effects that deletions confer upon a structure-function of FLuc, highlighting how specific regions play a role in different characteristics of the protein. As such, in order to evaluate this method and whether the Ω loop variants are truly useful, these mutants shall be further characterised following purification.

Either *in vivo* colony screening or 96-well format screening should be applied depending on the characteristics being sought, and the 96-well format is advantageous when screening for a number of parameters simultaneously especially from large libraries. This will ensure that potentially useful mutants are sampled and that mutations in regions that appear to bestow useful characters are not overlooked.

4.5. Conclusion

In this chapter an alternative screening strategy was developed in order to address drawbacks of the *in vivo* colony screen and to accelerate the identification of useful mutants. In particular the alternative strategy focused on the development of a method able to quantify bioluminescent activity by multi-parameter 96-well assays.

Initial testing of WTFLuc, x2FLuc and x11FLuc controls provided evidence of its utility. With regards to the single amino acid deletion mutants, a trend initially observed throughout *in vivo* colony screening by ImageJ analysis was confirmed by this method, showing deletions within the N-terminal caused little effect on the properties of x11FLuc, whilst C- terminal deletions caused a reduction in activity. Deletions occurring between L172- T191 and D520-L526 largely reduced activity. Whilst deletions occurring within T352-F368 and D375-R387 caused variation in activity, and some of the highest activity imparted by deletions occurred within these regions.

Deletions occurring within the Ω loop confer interesting properties upon the protein, including enhancements in overall activity, specific activity and altered kinetics and as such it has been concluded that deletions occurring within the Ω loop shall be taken forward not only to characterise this further but to determine the utility of the 96-well format.

Chapter 5

Biochemical Characterisation of Single Amino Acid Deletion within the Omega Loop of Luciferase

5.1. Chapter Summary

Previous Chapters showed that multiple phenotypes (enhanced activities and bioluminescence shifts) were obtained by SDD within the Ω loop structure of x11FLuc, however, all observations were made using colony level and 96-well format cell lysate screening assays. Characterisation of mutant FLucs following purification provides the most meaningful results, for example, the basic biochemical properties of the enzymes. In addition to this, it is possible to correlate the 96-well format screening with that at the level of protein. Within this chapter x11FLuc single amino acid deletion mutants ΔP353, ΔP359, ΔG360, ΔA361, ΔV362 and ΔG363, derived from the Ω loop, were purified and properties such as bioluminescence spectra, kinetics, specific activity and resistance to thermal inactivation and pH tolerance were analysed. Analyses revealed interesting properties, highlighting the importance of the Ω loop in the activity of x11FLuc. In addition, results correlated well to the 96-well format mutant fingerprinting method, further confirming the usefulness of the method.

5.2. Introduction

The number of applications in which Fluc may be utilised is extensive and mutants exhibiting enhanced characteristics are desirable within many fields (Ronaghi *et al.*, 2001; Branchini *et al*., 2009; Taurianen *et al.*, 2000). In Chapters 3 and 4, it was identified that single amino acid deletions in the Ω loop of FLuc cause interesting and potentially useful phenotypes, such as colour shifts and improved specific activity.

The Ω loop structure was first described in 1986 and subsequently characterised by Fetrow *et al*., (1995) and since then, the importance of this region has become clearer. In the recent past, protein engineering focusing on modulation of Ω loop structures

within other model proteins has proved successful. Guntas *et al* (2012) utilised a novel directed evolution engineering strategy termed circular permutation which involved random intramolecular relocation of the N- and C-termini within TEM-1 β-lactamase. This study isolated two permutated mutants within the Ω loop adjacent to the active site which conferred an increase in resistance to cefotaxime. In addition to this, triple mutants W165Y/E166Y/P167G occurring within the Ω loop indicated that substrate specificity could be greatly altered, increasing the hydrolytic activity for oxyimino-cephalosporin and ceftazidime whilst activity for other β-lactams was decreased (Stojanoski *et al*, 2015). Structural analysis of the triple mutant revealed that there was a large conformational change in the Ω loop creating space for the ceftazidime side chain. Their results indicated that the plasticity of the active site Ω loop facilitated the evolution of the enzyme specificity and mechanism (Stojanoski *et al*, 2015).

Substitutions with the Ω loop of *Ppy* FLuc (E354R/D357Y and E354I/D357Y) have previously been identified to improve thermostability and affect emission colour (see Chapter 1), however, to date, the significance of this loop in FLuc has not been identified. Tafreshi *et al* (2007), studied region (see Chapter 3) by insertion mutagenesis and observed both colour and thermostability phenotypes. These changes were attributed to a significant conformational change that had occurred within the loop bringing about a new ionic interaction affecting the polarity of emitter site.

The aim of this Chapter is to examine the effect of deletions in the Ω loop of x11FLuc (ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363). ΔT352 and ΔK358 were omitted within this study due to red shifted bioluminescence and reduced activity respectively. For this, mutants were overexpressed, purified and characterised in terms of bioluminescence spectra, kinetics, activity, thermostability and pH stability compared to WTFLuc, x2FLuc and x11FLuc.

5.3. Results and Discussion

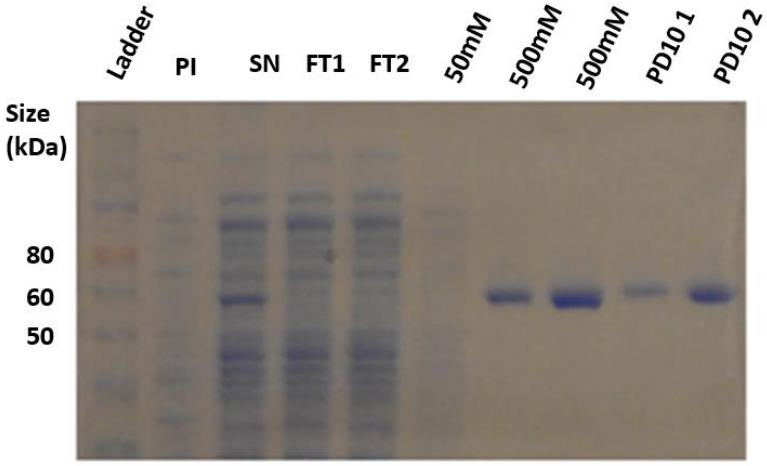
5.3.1. Overexpression and purification of single amino acid deletion variants

Single amino acid deletion variants containing a N-terminal x10 His-tag were overexpressed within *E.coli* BL21 cells and purified utilising affinity-based method using nickel-nitrilotriacetic acid (Ni-NTA) resin, as previously described (Law *et al.*, 2006). Proteins were overexpressed as described in Chapter 2. Fractions were analysed by SDS-PAGE (Figure 5.1.), the results of which were concurrent with enrichment calculations (Table 5.1.). A small amount of protein was lost in flow-through 1 and 2 of the supernatent, as well as the 50mM imidazole (IMD) wash, however the majority of expressed protein was isolated and eluted using fractions of 500mM IMD. Coomassie stained SDS-PAGE gels of each mutant showed no additional banding and the proteins were >90% pure. Eluted fractions displaying high activity were immediately desalted using PD10 size exclusion columns (GE Healthcare, WI, USA) into storage buffer (Chapter 2) and proteins were stored at -80^{o}C. The concentration of each PD10 desalted purified fraction was determined via Bradford assay (Table 5.2). Concentrations achieved throughout purification varied between 0.413mg/ml to 1.35mg/ml and in order to ensure these calculated concentrations were correct, all concentration normalised (0.188mg/ml) proteins were analysed by SDS-PAGE (Figure 5.2.). This showed the quantification of proteins was accurate, allowing further comparison of enzymes.

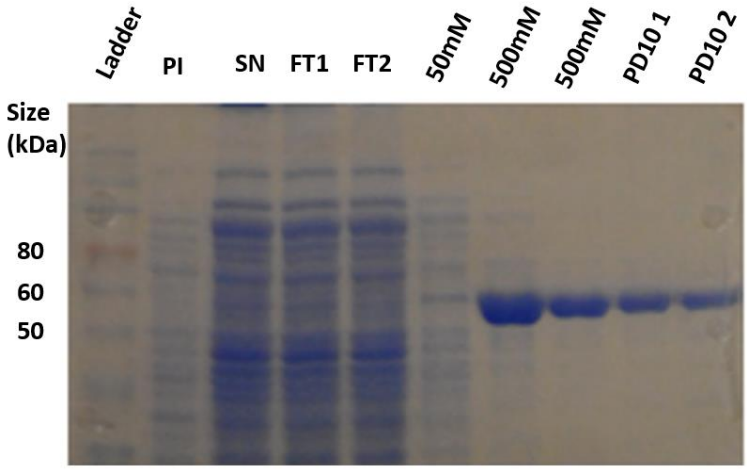
Figure 5.2 displayed an difference in molecular weight for the WTFLuc in comparison to other variants. Such "gel shifting" may be attributed to the tertiary structural changes that may have occurred within the protein as a result of the additional mutations comprising the x11FLuc since protein tertiary structure may affect both detergent-loading levels and polypeptide-SDS-PAGE migration rates (Rath *et al.,* 2008).

Figure 5.1. SDS-PAGE of Protein Purification

a.



b.



Example SDS-PAGE of control and mutant purification process. Purification of **a.** WTFLuc and **b.** x11FLuc. PI: Pre-induction sample, SN: supernatant, FT1: flow-through 1, FT2: flow-through 2, 50mM: 50mM IMD, 500mM: 500mM IMD, PD10: Desalted fraction.

Table 5.1. Representation of Activity in Purified Fractions

Fraction	**Vol (ml)**	**Average Imax(RLU)**	**Total Activity (RLU)**	**Enrichment (% activity of crude extract)**	**Enrichment in PD10 fractions (% of 500mM IMD fractions)**
S/N	20	3.99E+05	7.97E+06	100	
FT1	20	1.07E+04	2.13E+05	2.673662207	
FT2	20	3.43E+03	6.87E+04	0.861120401	
50IMD	10	1.46E+04	1.46E+05	1.826923077	
500IMD 1	2.5	1.35E+06	3.38E+06	42.33904682	
500IMD 2	2.5	1.27E+06	3.18E+06	39.88294314	
PD10 1	3.5	1.26E+06	4.41E+06	55.33862876	55.33862876
PD10 2	3.5	1.30E+06	4.55E+06	57.06521739	57.06521739

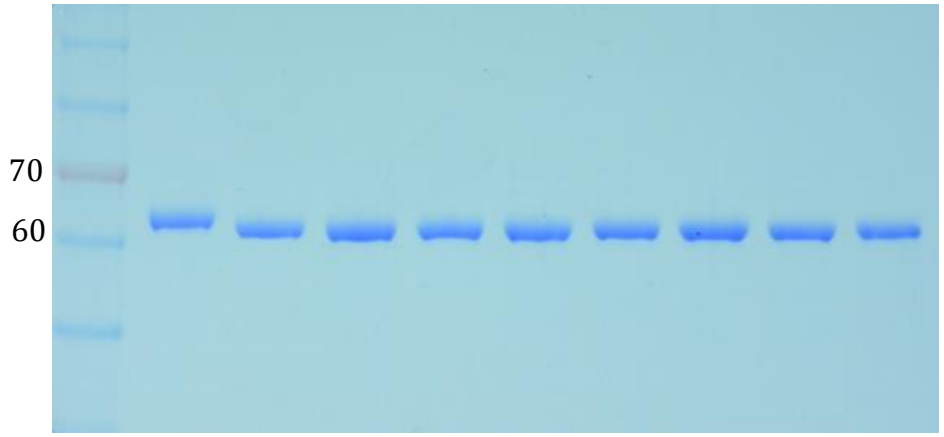
Summary of activity in purified fractions from WT. Key S/N: supernatant, FT: flow through, 50IMD: 50mM IMD, 500IMD: 500mM IMD. The Imax of 1µl of each fraction was assayed in saturating conditions using 500mM D-LH_2 and 1mM ATP. Vol. pertains to the fraction volume. Average Imax is the average observed Imax. Enrichment is the percentage activity of the crude extract in any one fraction in comparison to the supernatant. Enrichment of PD10 fractions considers only those desalted fractions against their initial non-desalted fractions. All activities were measured as flash heights.

Table 5.2. Summary of Average Protein Concentrations

Purified and desalted variant	WTFLuc	x2FLuc	x11FLuc	ΔP353	ΔP359	Δ G360	Δ A361	Δ V362	Δ G363
Average Concentration (mg/ml)	0.944	1.208	1.350	0.556	0.646	0.773	0.569	0.531	0.413
Average Concentration (□M)	15.14	19.37	21.65	8.92	10.35	12.38	9.13	8.52	6.63

Summary of Average Protein Concentration. The concentration of each PD10 desalted purified fraction as determined via Bradford assay. Concentrations presented both as mg/ml and corresponding μM.

Figure 5.2. SDS-PAGE of Normalisation of Proteins



SDS PAGE analysis for protein quantification. Proteins normalized to 0.1875mg/ml and applied to gel. Band size and density compared. All Lucs were mixed 3:1 in 4x protein sample buffer and loaded in the wells. Lanes 1. Molecular weight marker; 2. WT, 3. x2, 4. x11FLuc; 5. ΔP353, 6.ΔP359, 7.ΔG360, 8.ΔA361, 9.ΔV362, 10. ΔG363.

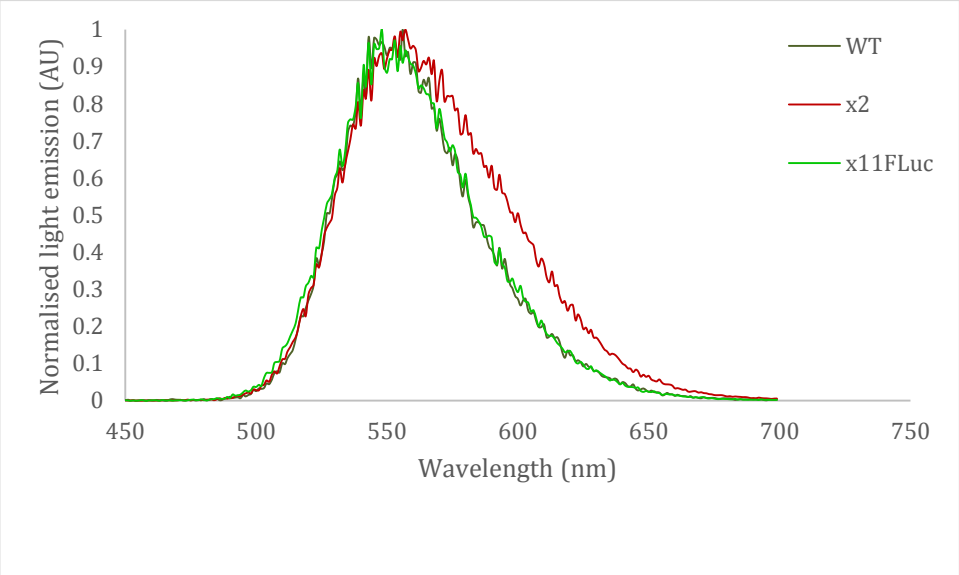
5.3.2. Bioluminescence spectra of x11 Deletion Mutants

To begin with, the bioluminescence spectra of purified WTFLuc, x2FLuc and x11FLuc control enzymes were measured first using a Varian Cary Eclipse fluorimeter in the presence of saturating conditions of LH_2 and ATP in TEM buffer at pH 7.8. Spectra were corrected for changes in the sensitivity of photomultiplier tube (PMT) to differences in the wavelengths of light emitted, as described previously (Law *et al*., 2006). The spectra of WTFLuc, x2FLuc and x11FLuc at pH, 7.8 peaked at ca. 550nm (Figure 5.3.). The bioluminescence spectra of single amino acid deletion variants in the Ω loop (ΔP353, ΔP359, ΔG360, ΔA361 and ΔV362) displayed □max similar to x11FLuc at pH 7.8 (λmax = ca. 555nm) (Figure 5.4.). However, variant ΔG363 showed a small increase in bandwidth toward the red region, indicating a direct or indirect specific change in the emitter.

5.3.3. pH dependence of bioluminescent spectra of x11FLuc single amino acid deletion variants

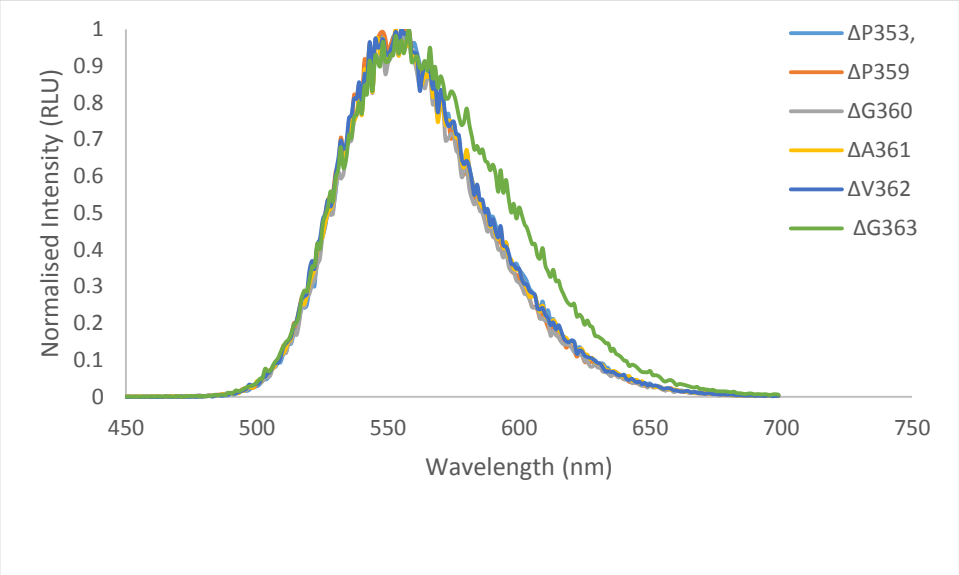
The bioluminescence spectra of single amino acid deletion variants were measured under various pH conditions (Figure 5.5-5.7 and Table 5.3). Ω loops variants, ΔP353, ΔP359, ΔG360, ΔA361 displayed similar spectra to x11FLuc (λmax ca.555nm) between pH 6.3 - 8.8, showing that deletions had little effect on the high degree of spectral stability of x11FLuc. However, Ω loop variants, ΔV362, ΔG363 did not follow this trend and bathochromic shifted at low or high pH, in a similar manner to x2FLuc. It is interesting that ΔG363 displayed spectral broadening more so under alkali conditions than under acidic pH, as is typically observed. Therefore, the conformation of the loop comprising residues V362 and G363 appears important for the emission colour.

Figure 5.3. Bioluminescence spectra of WTFLuc, x2FLuc and x11FLuc



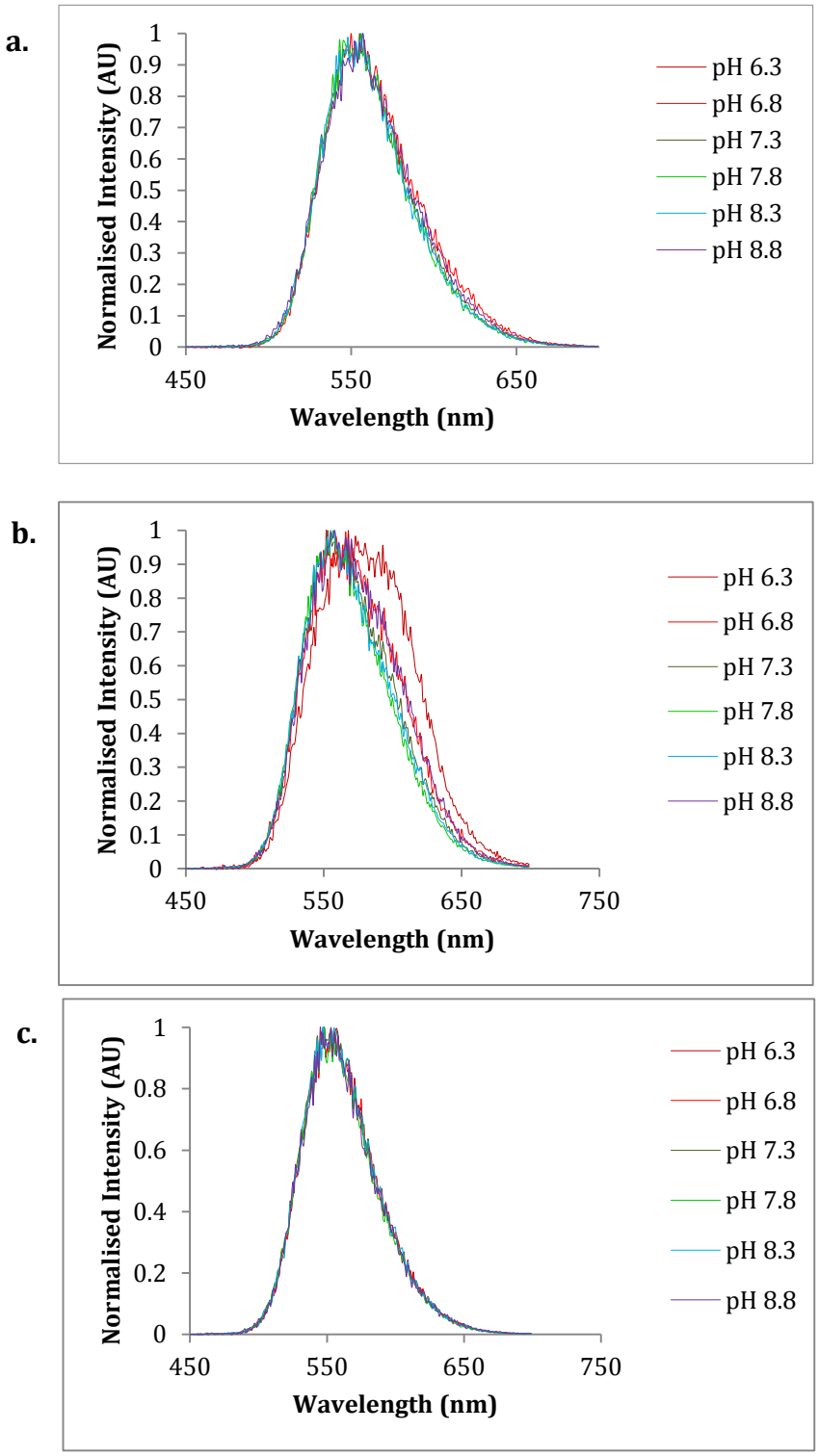
Baseline-corrected and normalised bioluminescence spectra of WTFLuc, x2FLuc and x11FLuc with LH_2 at pH 7.8. 0.5µM of each enzyme added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured after 15 seconds at RT.

Figure 5.4. Bioluminescence spectra of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363



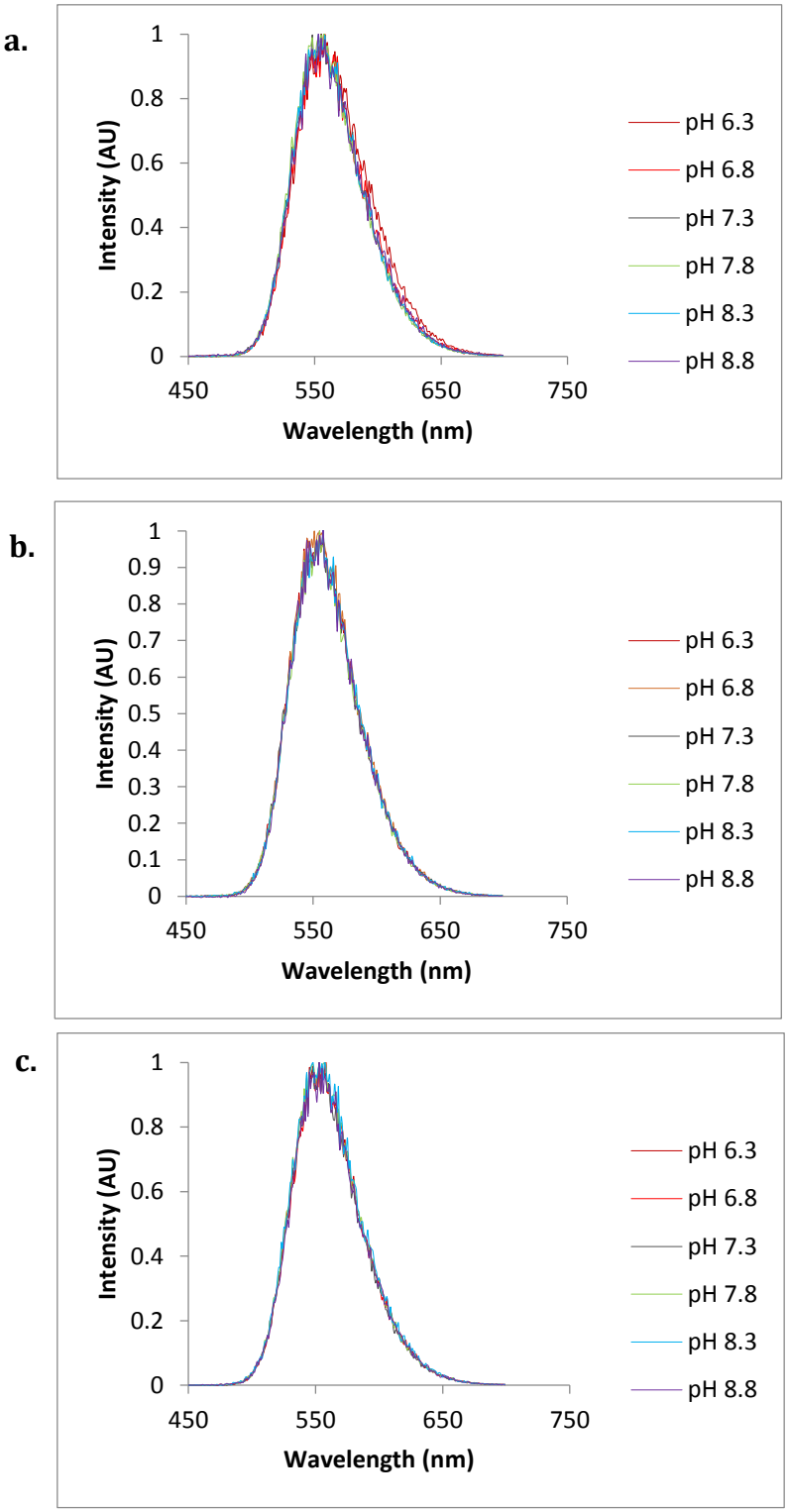
Baseline-corrected and normalised bioluminescence spectra of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 with LH_2 at pH 7.8. 0.5µM of each enzyme added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured after 15 seconds at RT.

5.5. pH dependence of bioluminescence spectra of WTFLuc, x2FLuc and x11FLuc



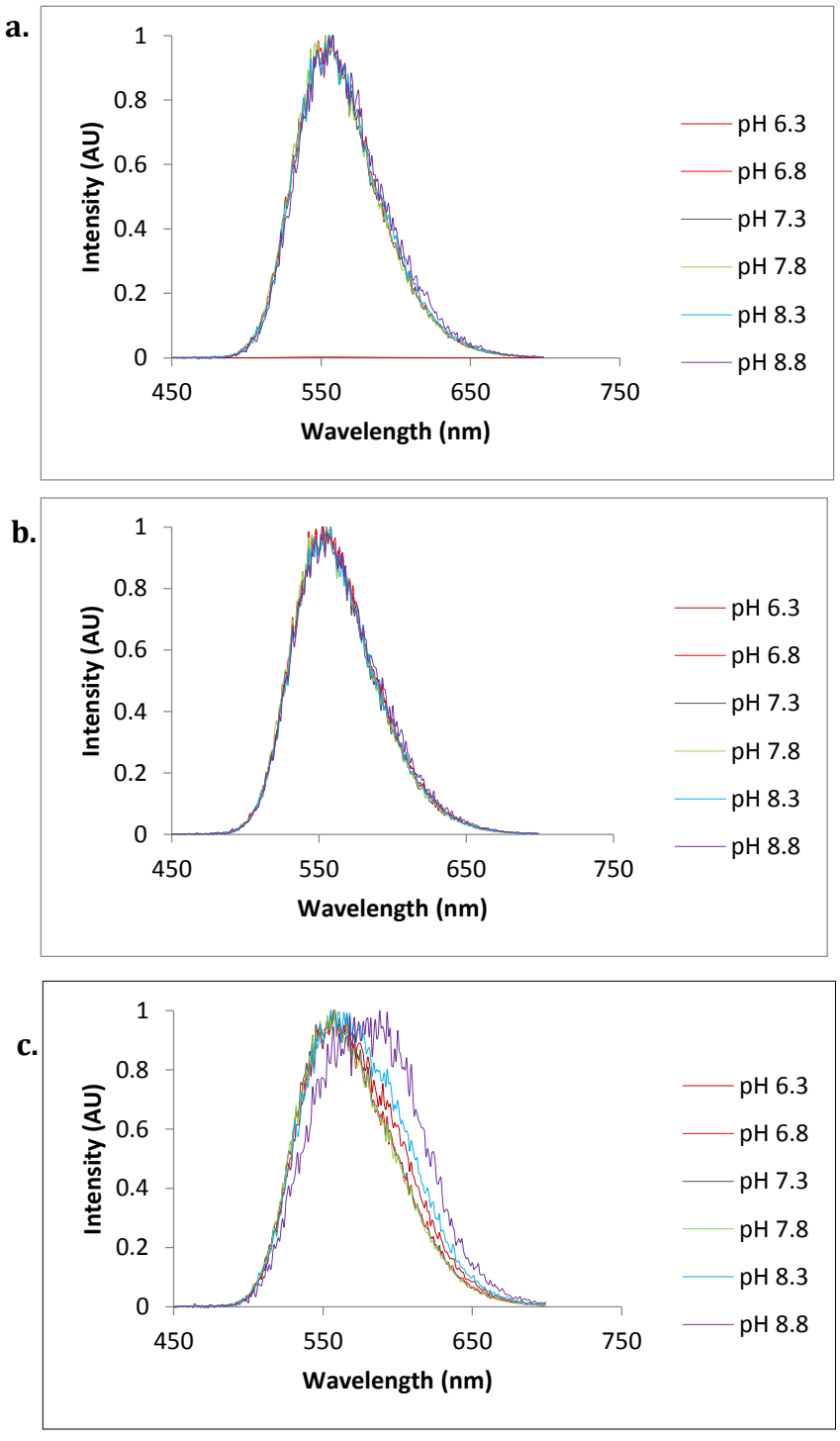
Baseline corrected and normalised bioluminescence spectra of WTFLuc, x2FLuc and x11FLuc at varying pH with LH_2. a) WTFLuc b) x2FLuc c) x11FLuc assayed with 0.5µM of each enzyme added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured after 15 seconds at RT.

5.6. pH-dependence of bioluminescent spectra of ΔP353, ΔP359, ΔG360



Baseline corrected and normalised bioluminescence spectra of ΔP353, ΔP359, ΔG360 at varying pH with LH_2. a) ΔP353, b)ΔP359, c)ΔG360 assayed as details in Figure 5.4.

Figure 5.7. pH-dependence of bioluminescence spectra of ΔA361, ΔV362, ΔG363



Baseline corrected and normalised bioluminescence spectra of ΔA361, ΔV362, ΔG363 at varying pH with LH_2. a)ΔA361, b)ΔV362, c)ΔG363. Details as in Figure 5.4.

Table 5.3. Summary of pH-Dependence of Bioluminescence Spectra

	Spectra											
	pH 6.3		pH 6.8		pH 7.3		pH7.8		pH 8.3		pH 8.8	
	Λmax	FWHM	Λmax	FWHM	λmax	FWHM	λmax	FWHM	λmax	FWHM	Λmax	FWHM
WTFLuc	N/A	N/A	553	61	554	57	553	58	554	59	554	63
x2FLuc	575	90	564	85	556	75	555	69	555	74	561	83
x11FLuc	554	61	554	62	554	62	554	62	554	62	554	62
ΔP353	558	64	555	64	555	61	555	60	554	66	554	66
ΔP359	551	62	554	60	553	58	554	61	554	59	554	64
ΔG360	554	62	554	61	554	62	554	61	554	64	554	64
ΔA361	555	61	555	61	554	62	554	63	555	60	555	60
ΔV362	553	63	554	64	555	60	555	61	555	60	555	63
ΔG363	563	81	558	73	558	75	558	75	565	86	580	90

Summary of characteristics of bioluminescence spectra exhibited by controls and deletions mutants at variable pH (between 6.3-8.8)

.

5.3.4. Michaelis-Menten Kinetic Characterization of Deletion Mutants Compared to WTFLuc, x2FLuc and x11FLuc

Since the bioluminescence reaction displays flash kinetics, conventional methodology may not be employed in order to determine the kinetic constants. However, the peak intensity (I_{max}) is a representation of the pre-steady state of maximal light intensity at a given substrate concentration and occurs at one turnover of the enzyme (Brovko *et al*, 1994). Use of I_{max} to determine kinetic constants obeys Michaelis-Menten kinetics for the reaction prior to other complicating factors such as end product inhibitors (Ugarova, 1989). To determine kinetic constants by luminometry, enzymes were exposed to varying concentrations of the substrate in question and other components were injected at saturating concentrations. The resulting triplicate I_{max} values for each concentration were plotted according a rearranged version of the Michaelis-Menten plot, termed the Hanes-Woolf plot. Within this investigation, the Hanes-Woolf plot has been utilised in favour of the Lineweaver-Burke plot since it better copes with experimental error (Athel-Cornish Bowden, 1999). Regression analysis of the Hanes-Woolf plot (substrate concentration (S) plotted against S/v, wherein v is the initial rate) allows determination of constants K_M (the Michaelis-Menten constant) and the V_{max} (maximal reaction velocity), with which it is possible to characterise the different mutant FLucs. The catalytic constant, termed k_{cat}, is further calculated from the V_{max} value and represents the turnover of the enzyme, whilst the ratio k_{cat}/K_M is a useful measure of overall catalytic efficiency. Substrate ranges used were 10x above and below the approximate K_M for enzymes, and as such concentrations of LH_2 were diluted such that the final working concentration within the reaction mix included 0.1µM-200µM and concentrations of ATP were diluted such that the final working concentration within the reaction mix included 0.1µM-1000µM.

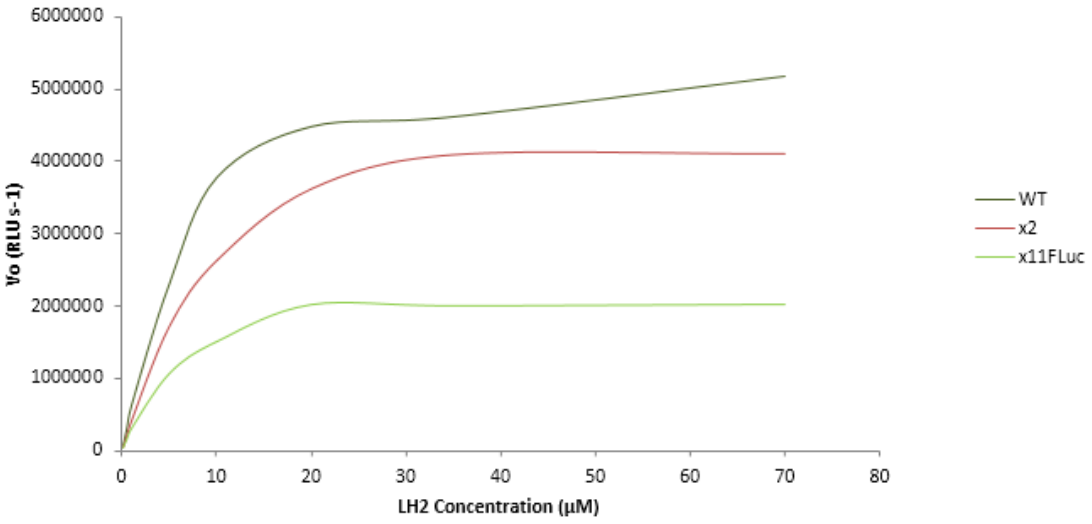
Kinetic parameters determined via this method were similar to those previously identified for the enzyme controls (Figure 5.8-5.11 and Table 5.4). In terms of K_m values for LH_2, the value for WTFLuc was 25µM, which was similar to that noted in previous studies (approximately 15µM) (Branchini *et al*., 1998; Tisi *et al*., 2002; Law *et al*., 2006). x2FLuc displayed a K_M for LH_2 of 10µM, which was very similar to that which is reported in the literature. The LH_2 K_M value derived for x11FLuc was 5µM, again, similar to previous studies (Jathoul *et al*., 2012). K_M values for ATP were as

follows: WTFLuc: 125µM was calculated at 125µM, x2FLuc: 42.5µM, and, x11FLuc: 75µM, which are in broad agreement with the literature.

The single amino acid deletion variants had similar K_M values to the parental x11FLuc (Figure 5.12-5.15, Table 5.5). ΔP353, ΔA361, ΔV362 and ΔG363 exhibited similar or slightly higher K_M than x11FLuc, of between 6µM - 11µM. However, deletions of ΔP359 and ΔG360 appeared to lower the K_M compared to x11FLuc, to 2µM and 3µM for ΔP359 and ΔG360 respectively. This indicates that the conformation of this region is important for the affinity of the enzyme toward LH_2. However, since on the whole very small changes in K_M for LH_2 were seen, it appears that the other single amino acid deletions in loops of x11FLuc do not have a significant effect on the affinity with LH_2. In regards to the K_M of enzymes for ATP, ΔP353 and ΔG360 exhibited similar values to x11FLuc of 66.66µM and 58.33µM, respectively. However, the K_M for ATP of mutant ΔP359 was increased to 100µM, whilst the other single amino acid deletion mutants ΔA361 (147µM), ΔV362 (150µM), ΔG363 (188µM) displayed a gradual increased K_M, which appeared to correlate with the position of the deletion with respect to its distance from the C-terminal of the Ω loop. The variation in the K_M for ATP was far more marked than variations in the K_M for LH_2 which indicates that the Ω loop plays a more important role in modulating the affinity for ATP.

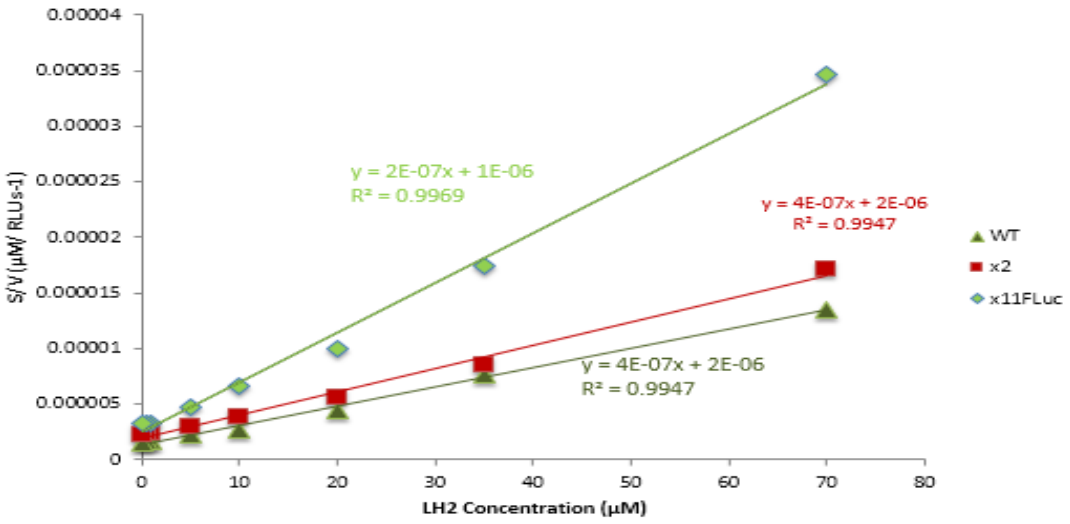
Statistical analysis was conducted to determine the significance of the differences observed (Figure 5.16-5.17). A statistically significant difference between WT, x2 and x11FLuc regarding KM Luciferin was determined by a one-way ANOVA, $F(2,3) = 67$, $p = 0.003$ however there was no statistically significant difference between WT, x2 and x11FLuc regarding KM ATP ($F(2,3) = 5.97$, $p = 0.09$). The Tukey method was further employed which identified that the WT KM Luciferin was significantly different from x2 and x11FLuc. Of the Ω loop variants, one way ANOVA determined there was a statistically significant difference regarding both KM Luciferin ($F(6,8) = 10.58$, $p = 0.002$) and KM ATP ($F(6,10) = 13.58$, $p = 0.00$). The Tukey method identified that there was a significant difference in the KM Luciferin between x11FLuc and ΔA361 and ΔP359 whilst a significant difference in the KM ATP was found between x11FLuc and ΔG363.

Figure 5.8. Michaelis-Menten plots of WTFLuc, x2FLuc and x11FLuc with LH_2



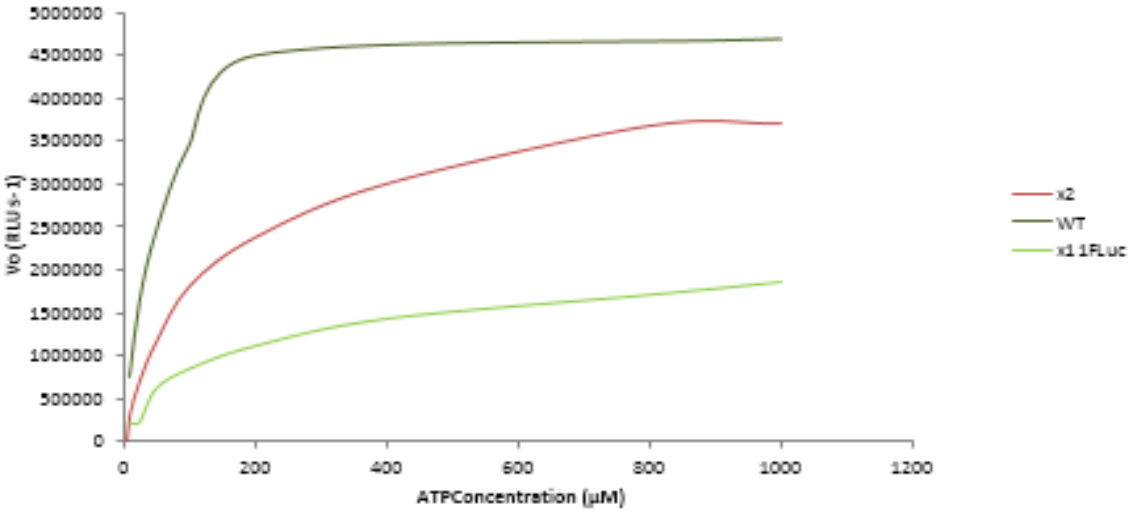
Comparative plots of initial velocities against substrate concentration. To measure LH_2 kinetic parameters, saturating ATP was injected at a final concentration of 1mM in the well prior to measurement of bioluminescence. Substrate ranges of LH_2 were ca. 0.1-10x the K_M concentrations. 2-4 independent experiments (each in triplicate) carried out and examples of each enzyme have been plotted.

Figure 5.9. Hanes-Woolf plot of WTFLuc, x2FLuc and x11FLuc with LH_2



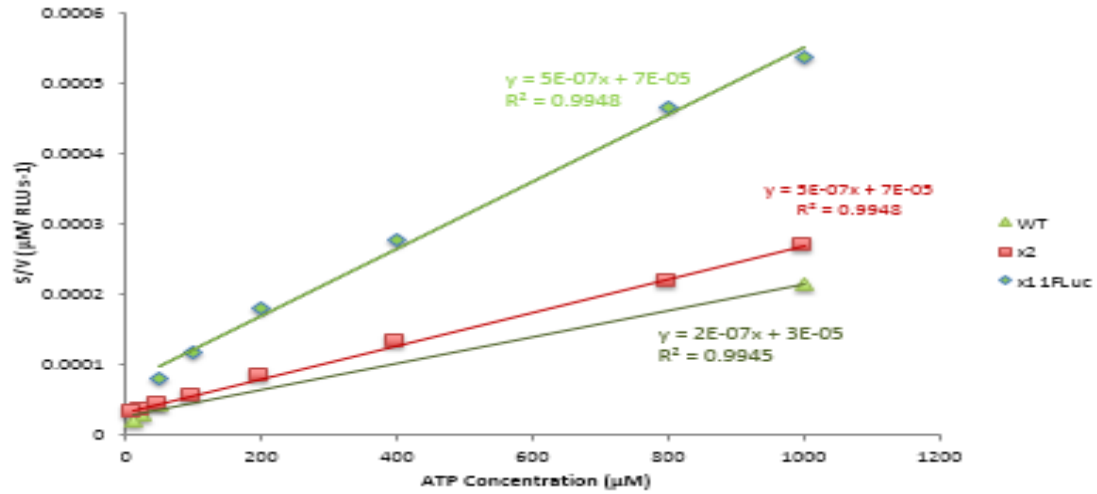
Hanes-Woolf plots to determine the Michaelis-Menten kinetic parameters. Plots derived from data indicated in Figure 5.7. by plotting [S]/v against [S], where substrate concentration and v is estimated initial rate at each concentration. Kinetic parameters calculated by linear regression of plots.

Figure 5.10. Michaelis-Menten plots of WTFLuc, x2FLuc and x11FLuc with ATP



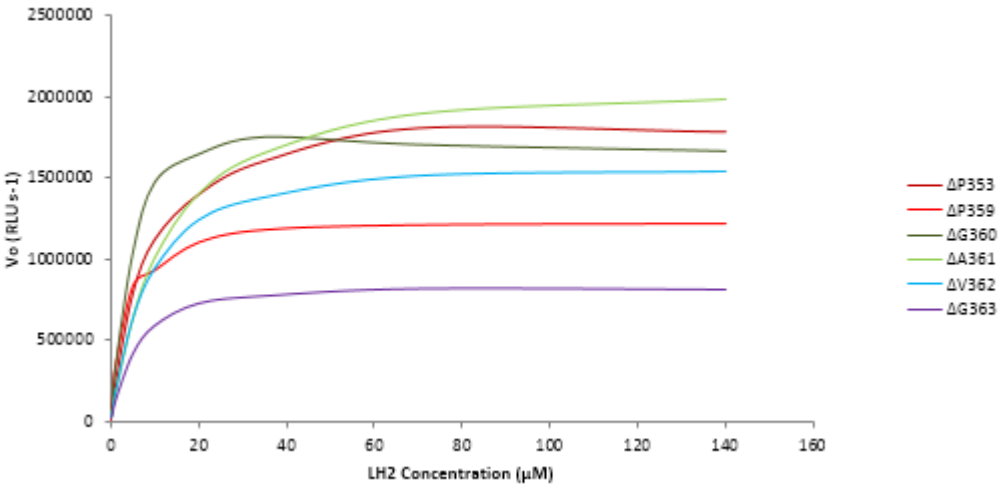
Comparative plots of initial velocities against substrate concentration. To measure ATP kinetic parameters, saturating LH_2 was injected at a final concentration of 500µM in the well prior to measurement of bioluminescence. Substrate ranges of ATP were ca. 0.1-10x the K_M concentrations. 2-4 independent experiments (each in triplicate) carried out and examples of each enzyme have been plotted.

Figure 5.11. Hanes-Woolf plot of WTFLuc, x2FLuc and x11FLuc with ATP



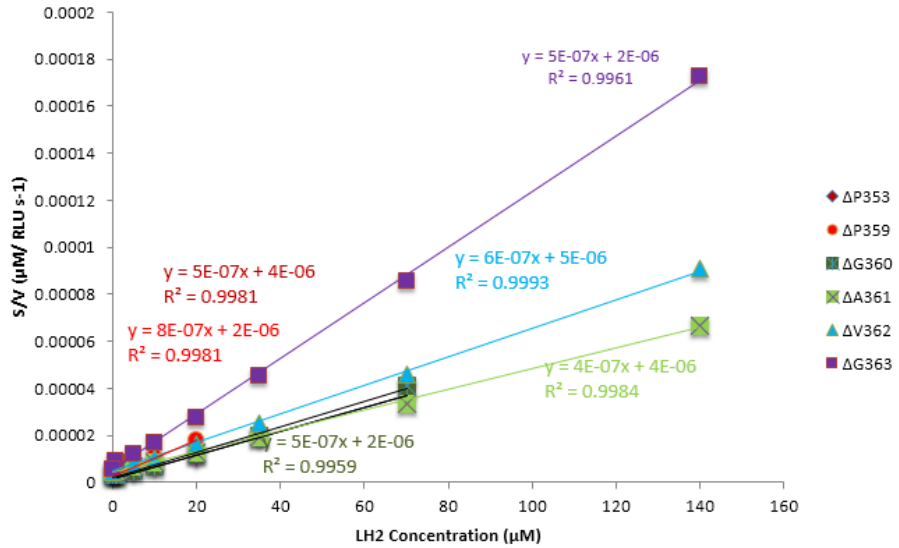
Hanes-Woolf plots to determine the Michaelis-Menten parameters. Plots derived from data indicated in Figure 5.10. Details as in Figure 5.9.

Figure 5.12. Michaelis-Menten plots of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 with LH_2



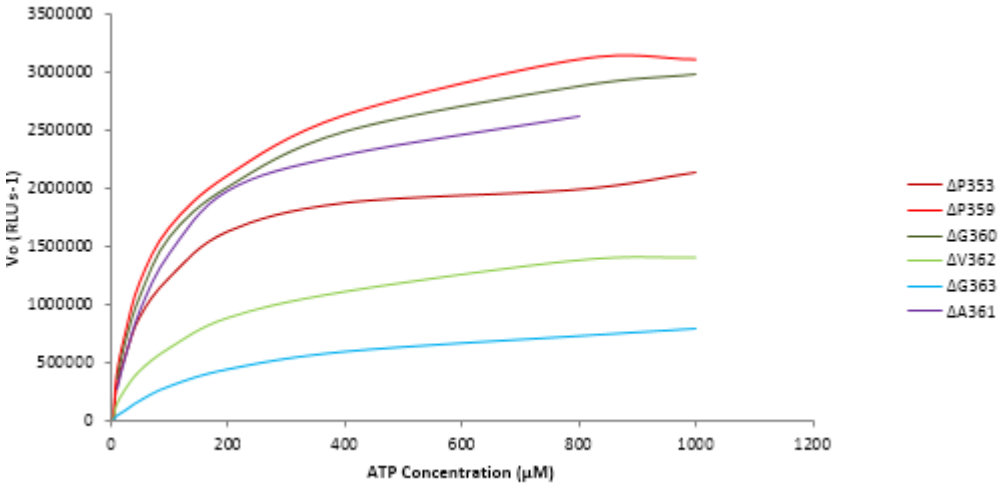
Comparative plots of initial velocities against substrate concentration. Details as in Figure 5.8.

Figure 5.13. Hanes-Woolf plot of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 with LH_2



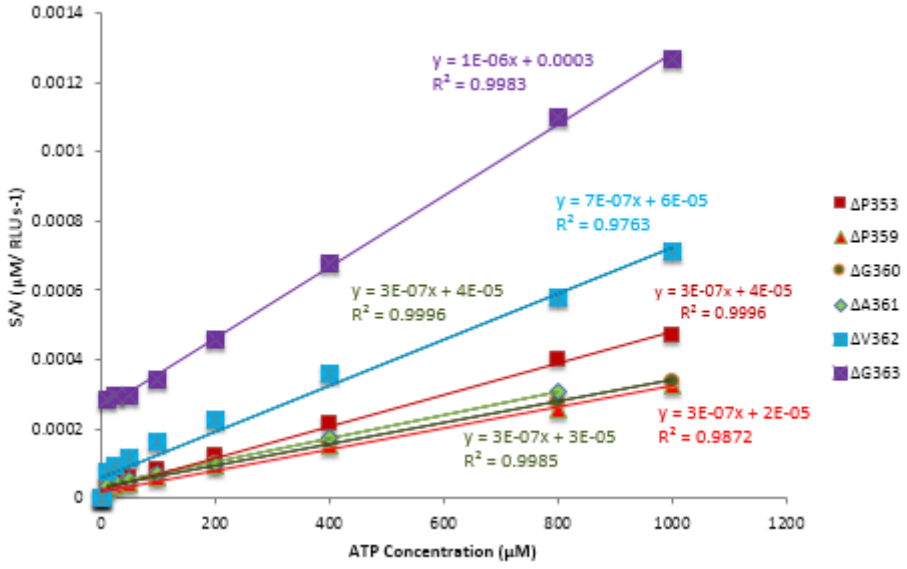
Hanes-Woolf plots to determine the Michaelis-Menten parameters. Details as in Figure 5.9

Figure 5.14. Michaelis-Menten plots of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 with ATP



Comparative plots of initial velocities against substrate concentration. Details as in Figure 5.10.

Figure 5.15. Hanes-Woolf plot of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 with ATP



Hanes-Woolf plots to determine the Michaelis-Menten parameters. Details as in Figure 5.11.

Table 5.4. Summary of kinetic parameters of WTFLuc, x2FLuc and x11FLuc with LH_2 and ATP

	LH_2			ATP		
	WTFLuc	x2FLuc	x11FLuc	WTFLuc	x2FLuc	x11FLuc
K_M(µM)	27.5±1.88	10	5	125±18.81	42.5±9.4	75.83±7.64
k_{cat} (RLU s-1)(x10^{16})	5	2.5	2	5	5	2.5
k_{cat}/K_M (RLU s-1 M-1	2.5x10^{15}	5x10^{15}	3.33x10^{15}	4x10^{14}	1.176x10^{15}	3.29x10^{14}

Summary of kinetic parameters for control enzymes. Standard errors shown. No error if not shown. k_{cat}/K_M derived from average K_M and k_{cat}.

Table 5.5. Summary of kinetic parameters of ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 with LH_2 and ATP

	LH_2						ATP					
	ΔP353	ΔP359	ΔG360	ΔA361	ΔV362	ΔG363	ΔP353	ΔP359	ΔG360	ΔA361	ΔV362	ΔG363
K_M(μM)	6±1.5	2.08±0.31	3.33±0.81	11±0.75	7.73±0.44	6	66.66±8.3	100	58.33±6.29	147.59±13.49	150±1.5	188.35±23.47
k_{cat} (RLU s-1)(x10^{16})	2	1.45±1.56	2	2.25±1.88	1.54±0.86	1	2	2	2	2.16±1.58	1.58±0.91	1
k_{cat}/K_M (RLU s-1 M-1)(x10^{15})	3.33	7	6	2.04	2	1.66	3	2	3.42	1.4	0.1	0.05

Summary of kinetic parameters for control enzymes. Standard errors shown. No error if not shown. k_{cat}/K_M derived from average K_M and k_{cat}.

Figure 5.16. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between K_M for Luciferin and ATP for Control Variants

a.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	2	558.33	279.167	67.00	0.003
Error	3	12.50	4.167		
Total	5	570.83			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
WT	2	27.50	A
x2	2	10.00	B
x11FLuc	2	5.000	B

b.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	2	6890	3444.9	5.97	0.090
Error	3	1731	576.9		
Total	5	8621			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
WT	2	125.0	A
x11FLuc	2	75.83	A
x2	2	42.5	A

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from WT, x2 and x11FLuc. a) A statistically significant difference between WT, x2 and x11FLuc regarding K_M Luciferin was determined by a one-way ANOVA, $F(2,3) = 67$, $p = 0.003$. b) A non statistically significant difference between WT, x2 and x11FLuc regarding K_M ATP was determined by a one-way ANOVA, $F(2,3) = 5.97$, $p = 0.09$. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. N=2.

Figure 5.17. One-Way ANOVA and Tukey HSD to Determine Statistically Significant Differences between K_M for Luciferin and ATP for Ω Loop Mutants

a.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	6	108.69	18.115	10.58	0.002
Error	8	13.69	1.712		
Total	14	122.38			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

ΔA361	2	11.00	A
ΔV362	2	7.720	A B
ΔP353	2	6.00	B C
ΔG363	2	6.000	B C
x11FLuc	**2**	**5.000**	**B C**
ΔG360	3	3.333	B C
ΔP359	2	2.080	C

b.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Variant	6	38549	6424.8	13.58	0.000
Error	10	4730	473.0		
Total	16	43279			

Tukey Pairwise Comparisons

Grouping Information Using the Tukey Method and 95% Confidence

Variant	N	Mean	Grouping
ΔG363	3	188.0	A
ΔV362	2	150.0	A B
ΔA361	3	147.59	A B
ΔP359	2	100.0	B C
x11FLuc	**2**	**75.83**	**B C**
ΔP353	3	66.67	C
ΔG360	2	58.34	C

Minitab session output displaying results of a One-way ANOVA and Tukey HSD test from data obtained from Ω loop deletion mutants. a) A statistically significant difference between ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 regarding K_M Luciferin was determined by a one-way ANOVA, $F(6,8) = 10.58$, $p = 0.002$. b) A statistically significant difference between ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 regarding K_M ATP was determined by a one-way ANOVA, $F(6,10) = 13.58$, $p = 0.00$. Grouping information using the Tukey method and 95% confidence. Means that do not share a letter are significantly different. x11FLuc control highlighted in bold. N=2.

5.3.5. pH-dependence of bioluminescent activity of WTFLuc, x2FLuc and x11FLucs

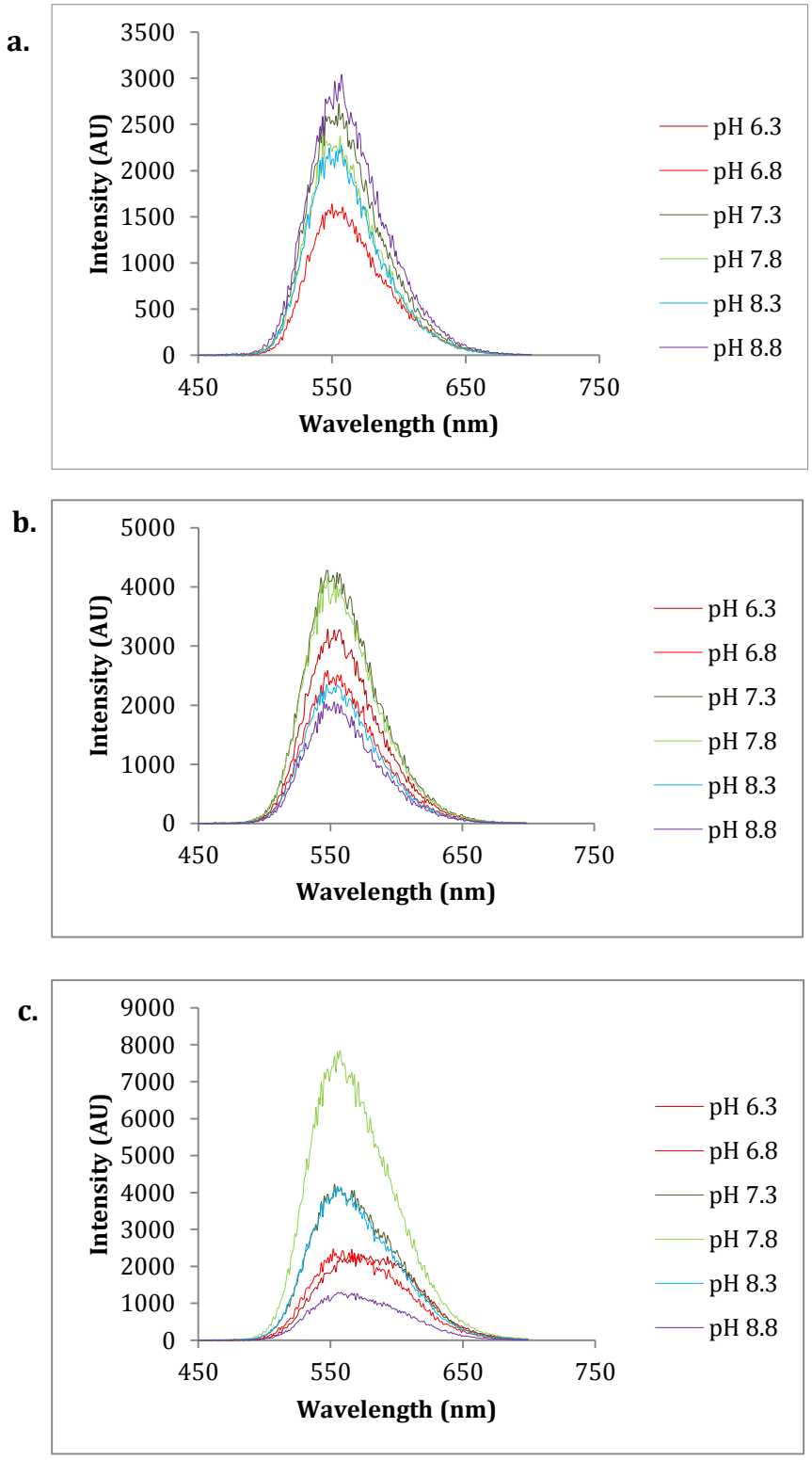
The activities of WTFLuc, x2FLuc and x11FLuc were measured at differing pH conditions by obtaining a bioluminescence emission spectra in the presence of saturating conditions of LH_2 and ATP at different pH values and integrating PMT-sensitivity corrected spectra (Figure 5.16. and Table 5.6.). WTFLuc was seen to retain up to 55% of activity at pH 6.8 (one unit lower than optimum), whilst activity diminished by up to 18% at pH 8.8 (one unit higher). It has previously been shown that x11FLuc better retains maximum activity across a wide pH range (Jathoul *et al*., 2012) and here x11FLuc was the most resistance to inactivation at pH 6.8 of all enzymes, retaining up to 63% of maximum activity.

Previous investigations into the bioluminescent spectra of x2FLuc have shown a bathochromic shift of this mutant (Jathoul *et al*., 2012). White *et al* (2002) reported that x2FLuc is resistant to red-shift at low pH, and displays □max of 550-560nm. However, here and in previous reports (Amit Jathoul, 2008, PhD thesis, University of Cambridge) x2FLuc displays a 9nm bathochromic shifted □max with increased bandwidth at all pH values compared to WTFLuc. At pH 6.8 broadening of peaks toward the red was observed along with a shift in peak emission. x2FLuc is utilised commercially by 3M within the hygiene monitoring (FLuc-based ATP detection) industry and this bathochromic shift could be a disadvantage for luminometric assays involving x2FLuc. PMTs commonly employed in luminometers work on the principle of the photoelectric effect, whereby the signal from colours of higher frequency are amplified more than those of lower frequency. As a result PMTs are less sensitive in detecting red light and red-shifted FLucs, for example in environments with low pH, would give lower signals. However the use of CCD-based devices would overcome this problem. x11FLuc however, was seen to resist bathochromic shift at low pH, with □max 554nm at all pH values.

The specific activities of Ω loop deletion variants (Figure 5.17., Figure 5.18. and Table 5.6.), showed high activity across a wide pH range compared to WTFLuc, illustrating that this attribute of x11FLuc had been retained despite the location of the loop deletion in the Ω loop. At their lowest activity values (least optimum pH) ΔP353, ΔP359,

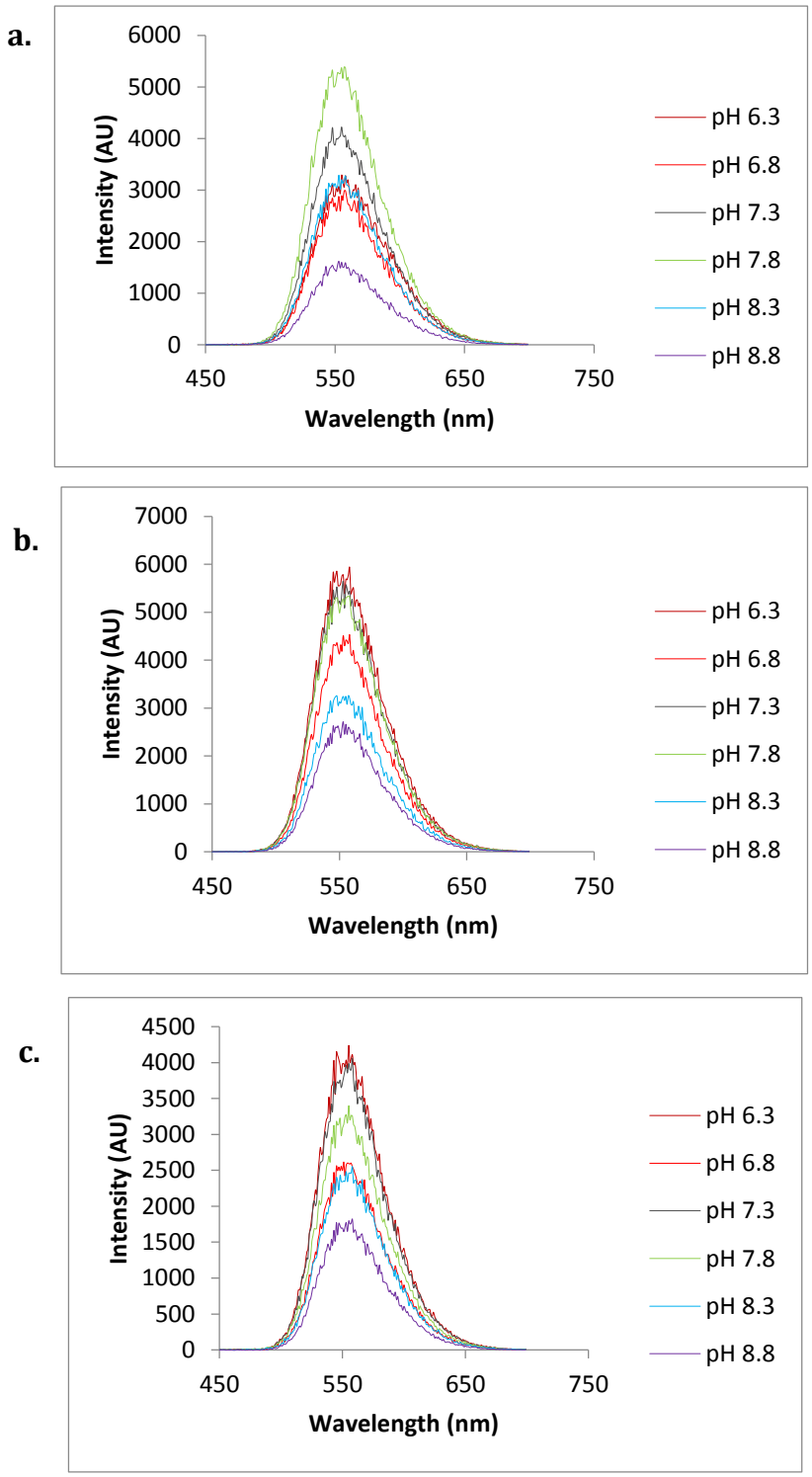
ΔG360, ΔA361, ΔV362, ΔG363 retained 30%, 45%, 42%, 24%, 31% and 19%, respectively.

Figure 5.16. pH-dependence of bioluminescent activities of WTFLuc, x2FLuc, x11FLuc



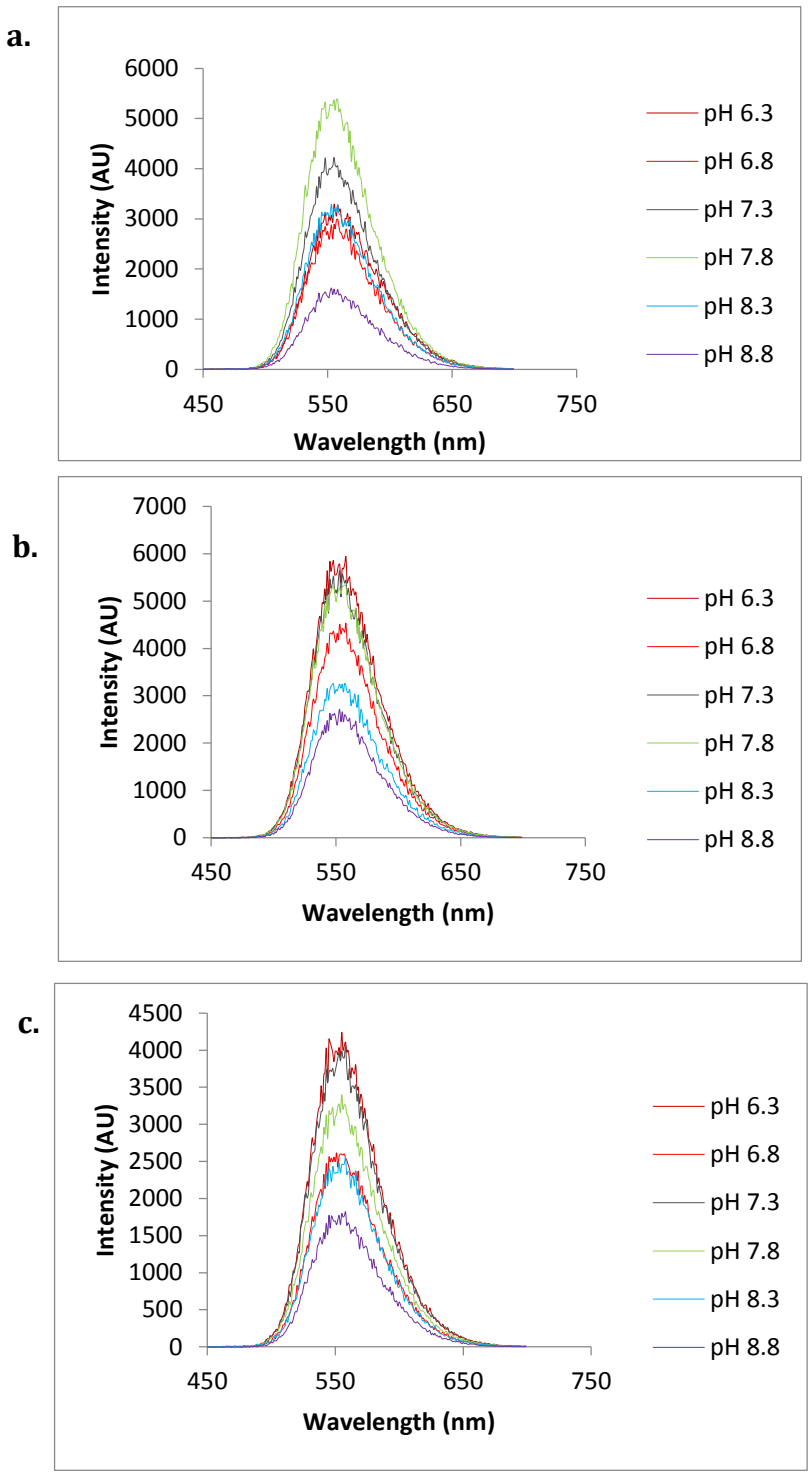
Baseline corrected bioluminescence spectra of WT, x2 and x11FLuc at varying pH with LH_2. a) WFLuc b) x2FLuc c) x11FLuc assayed with 0.5µM of each enzyme added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured after 15 seconds at RT.

Figure 5.17 pH-dependence of bioluminescent activities of ΔP353, ΔP359, ΔG360



Baseline corrected bioluminescence spectra of ΔP353, ΔP359, ΔG360 at varying pH with LH_2 a) ΔP353, b) ΔP359, c)ΔG360. Details as Figure 5.16.

Figure 5.18. pH-dependence on bioluminescent activities of ΔA361, ΔV362, ΔG363



Baseline corrected bioluminescence spectra of ΔA361, ΔV362, ΔG363 at varying pH with LH_2 a) ΔA361, b) ΔV362, c) ΔG363. Details as Figure 5.16.

Table 5.6 Summary of pH-dependence of Bioluminescence Activity

	pH 6.3		pH 6.8		pH 7.3		pH 7.8		pH 8.3		pH 8.8	
	Integrated Activity	Percentage of Max	Integrated Activity	Percentage of Max	Integrated Activity	Percentage of Max	Integrated Activity	Percentage of Max	Integrated Activity	Percentage of Max	Integrated Activity	Percentage of Max
WTFLuc	202693.6	0.36	300600.78	0.54	408550.1	0.73	552150.5	1	182143.13	0.32	108140.36	0.19
x2FLuc	242562.3	0.51	327497.61	0.69	417297.6	0.88	472071.5	1	273175.09	0.57	142918.40	0.3
x11FLuc	185090.1	0.63	229701.74	0.78	291910.1	1	201594.4	0.69	192435.16	0.65	77905.548	0.26
ΔP353	201418.2	0.34	395518.67	0.67	314331.5	0.53	588997.2	1	285974.87	0.48	146586.89	0.24
ΔP359	407138.7	0.89	387094.78	0.85	454709.7	1	366505.0	0.80	267744.29	0.65	179049.96	0.39
ΔG360	190460.8	0.69	273172.85	1	232842.2	0.85	231874.2	0.84	258144.28	0.94	102022.86	0.37
ΔA361	247034.4	0.64	317869.95	0.82	384749.5	1	368690.7	0.95	275712.65	0.71	87827.41	0.22
ΔV362	254004.9	0.57	42380.223	0.096	440395.6	1	384117.0	0.87	374242.84	0.84	111754.12	0.25
ΔG363	274232.8	0.64	409059.31	0.96	58839.54	0.13	424569.2	1	312691	0.73	107376.52	0.25

Summary of pH-dependence of bioluminescence activity. Details as in Figures 5.16–5.18.

5.3.6. Specific Activity of x2FLuc, x11FLuc and Single Amino Acid Deletion Variants Derived from Flash Kinetics by Luminometry

Assays were conducted under saturating conditions of the substrates and total bioluminescent output studied over a 250s period, integrating every 1s in order to compare differences between flash heights (Figure 5.19.) and overall bioluminescent output using a BMG Fluostar Optima (BMG Labtech, Ortenberg, Germany) (Table 5.7.). Interestingly, all mutants tested, apart from x11FLuc ΔP353 displayed higher specific activities than the x11FLuc. In order from lowest to highest, ΔP353 displayed an integrated light-based specific activity of 10195718 RLU (54%), whilst ΔP359, ΔV362 and ΔG363 displayed activities of 38758893 RLU, 30381748 RLU and 31124582 RLU (234%, 183% and 188% of x11FLuc), respectively. x11FLuc ΔG360 displayed 16572912 RLU (100% of x11FLuc) and ΔA361: 26156105 RLU (158% of x11FLuc). Therefore, deletions made in the Ω loop, aside from at position 353, promoted an increase in the specific activity compared to native x11FLuc when measuring light output over long periods (Figure 5.20.).

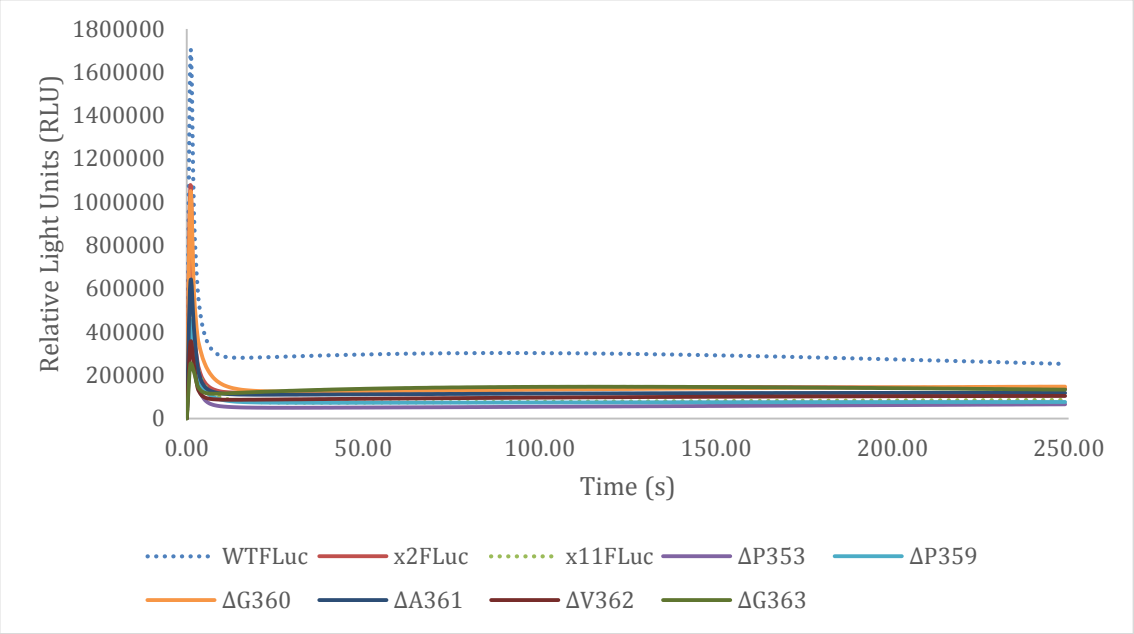
.

Table 5.7 Summary of Integrated Bioluminescence between 30-250s

Variant	**Integrated Bioluminescence between 30s-250s (RLU)**	**Percentage of x11Fluc (%)**
x11FLuc	16587278	100
ΔP353	9020908	54.3845
ΔP359	38758893	233.6664
ΔG360	24859369	149.8701
ΔA361	26156105	157.6877
ΔV362	30381748	183.1629
ΔG363	31124582	187.6413

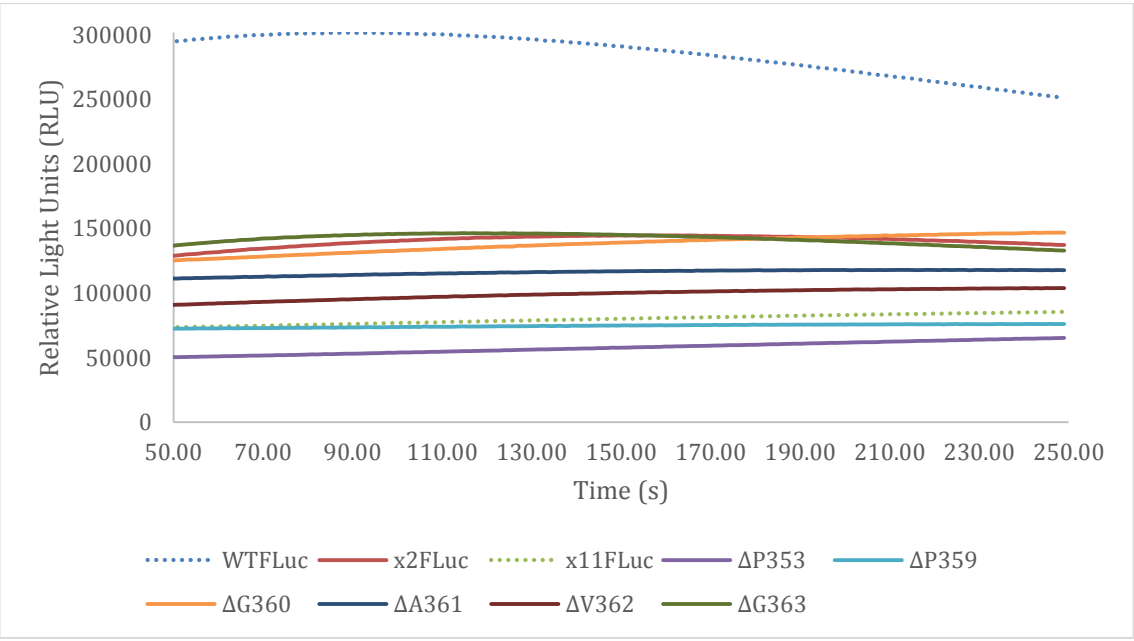
Table of integrated bioluminescence values measured by luminometery. Activity of x11FLuc is increased by deletional mutagenesis.

Figure 5.19. Flash Kinetics of FLucs and Mutants Over 250s



Flash kinetics of controls and mutants over time. Flash height of x11FLuc is increased by deletional mutagenesis.

Figure 5.20. Steady-State Kinetics of FLucs and Mutants Over 250s



Steady-state kinetics of controls and mutants over time after the flash. Specific activity of x11FLuc is increased by deletional mutagenesis.

5.3.7. Thermal Inactivation of Single Amino Acid Deletion Variants

There are two methods in which thermal inactivation of a protein may be studied, the conventional method, which involves the heating the enzyme to a specified temperature and subsequently cooling it on ice prior to taking measurements. Another method involves the measurement of activity whilst at the specified temperature. As part of this investigation a study was undertaken to characterise the degree to which the x2FLuc, x11FLuc and the Ω loop deletion enzymes performed under a range of temperatures using both these methods.

x2FLuc resisted thermal inactivation up to 40ºC (Figure 5.21.), however, above this temperature it is known to inactivate markedly, with 0% of initial activity remaining after 20 minutes at 44ºC (White *et al*., 1996). In agreement, here x2FLuc was observed to rapidly inactivate at 45ºC, retaining 22% of initial activity after 60 minutes and only 0.002% after 60 minutes at 50ºC.Throughout the investigation x11FLuc performed similar as described in the literature, retaining activity up to 93% of activity following 60 minutes incubation at 50ºC (Jathoul *et al*., 2012).

The Ω loop variants displayed variable resistance to thermal inactivation up to a temperature of 45ºC suggesting that deletions affect the thermostability of x11FLuc. However, x11FLuc Ω loop variants outperformed x2FLuc with regards to resistance to thermal inactivation. The largest difference between x11FLuc and deletion variants occurred at 50ºC over a 60 minute time interval (Figure 5.21.). At this temperature, all thex11FLuc Ω loop variants inactivated more than x11FLuc. x11FLuc P353 and ΔP359 displayed activity of 35% (as opposed to >90% for x11FLuc) and x11FLuc ΔV362 retained the highest level of activity of the Ω loop deletion mutants, of up to 72% of initial activity, whilst ΔG363 retained the lowest level (14%) of initial activity after 60 min at 50ºC.

Thermal inactivation of x2FLuc, x11FLuc and the Ω loop deletion variants were measured whilst enzymes were incubated at 3 temperatures: 30ºC, 35ºC and 42ºC (Table 5.8.), which are pertinent to the temperature range at which diagnostic ATP assays may be carried out. At these temperatures, activity amongst all Ω loop variants bar ΔG363 was increased in comparison to x2FLuc and x11FLuc. Of the temperatures studied ΔP353, ΔP359, ΔG360, ΔA361 were more stable, akin to x11FLuc, whilst ΔV362, ΔG363 were less stable, similar to x2FLuc, with maximal activity at 35ºC.

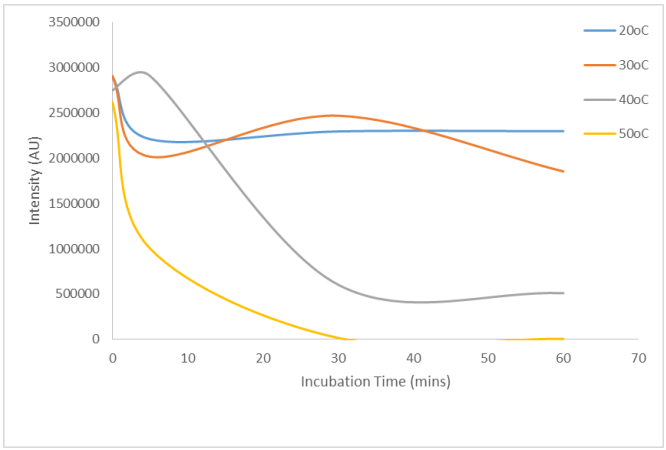
Table 5.8. Thermal Inactivation Assay Conducted at Temperature with x2, x11FLuc, ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363

	Temperature		
	30°C	35°C	42 °C
x2FLuc	386764.3	400024.7	395468
x11FLuc	440094.3	529217.3	559573.3
ΔP353	852537	940973	954703.7
ΔP359	554158.3	671513.3	689091.3
ΔG360	729496	858282	878304
ΔA361	597571.7	629020.3	634483.7
ΔV362	507992.7	561131.3	263037.3
ΔG363	232865	280963.3	17016

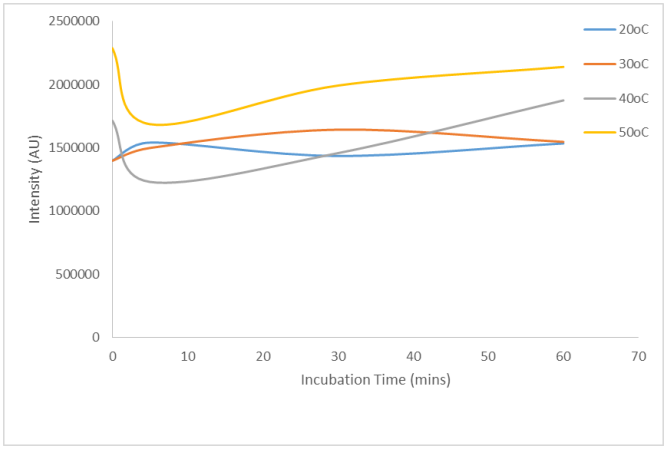
Bioluminescence activity measured after equilibration to temperature. Assays reflect real assay conditions to which FLucs may be subjected.

Figure 5.21. x2FLuc, x11FLuc and ΔP353bioluminescent activity measured at varying temperatures

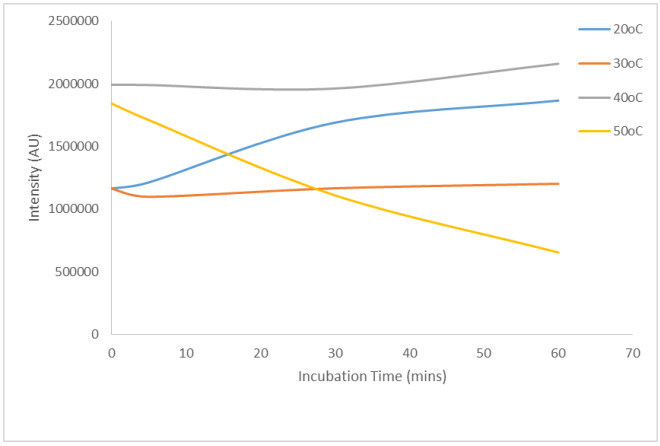
a.



b.



c.

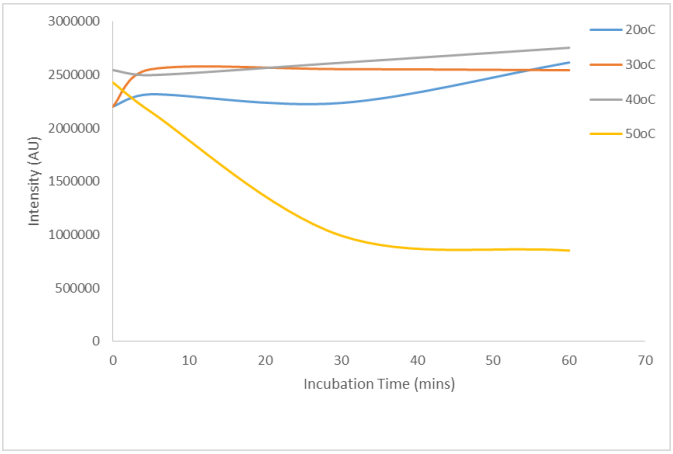


Bioluminescent activity displayed by variants following incubation at various temperatures. a) x2FLuc b) x11FLuc c)ΔP353 assayed with 0.5µM of each enzyme, pre incubated at different temperatures, added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured.

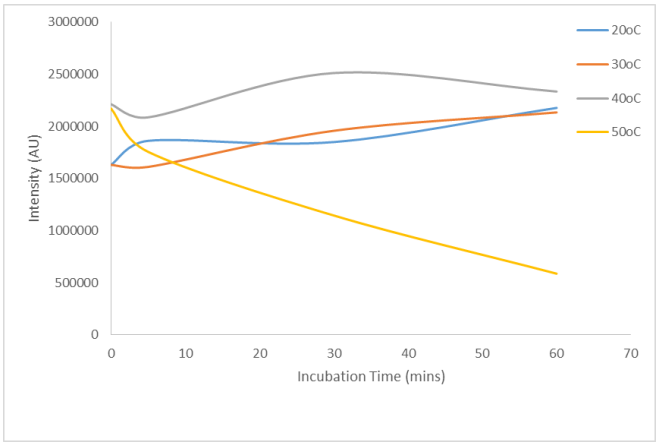
Figure 5.21. ΔP359, ΔG360 and ΔA361 bioluminescent activity measured at varying temperatures

.

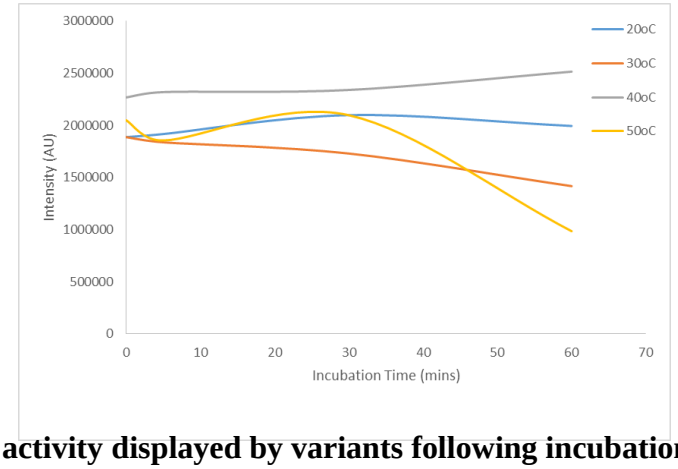
a



b.



c.

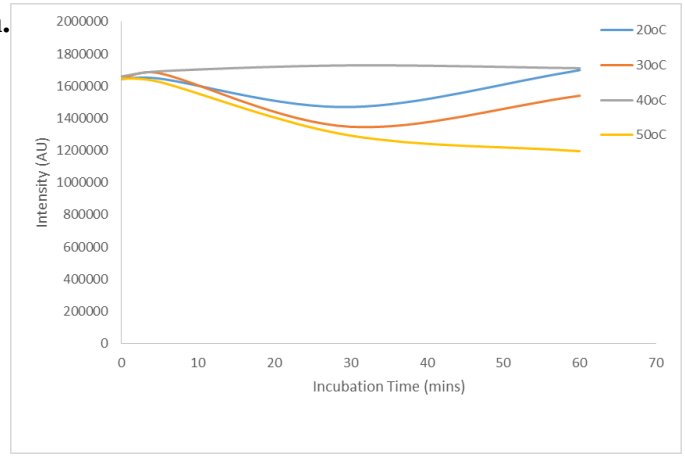


Bioluminescent activity displayed by variants following incubation at various temperatures. a) ΔP359 b) ΔG360 c) ΔA361 assayed with 0.5µM of each enzyme, pre-incubated at different temperatures, added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured.

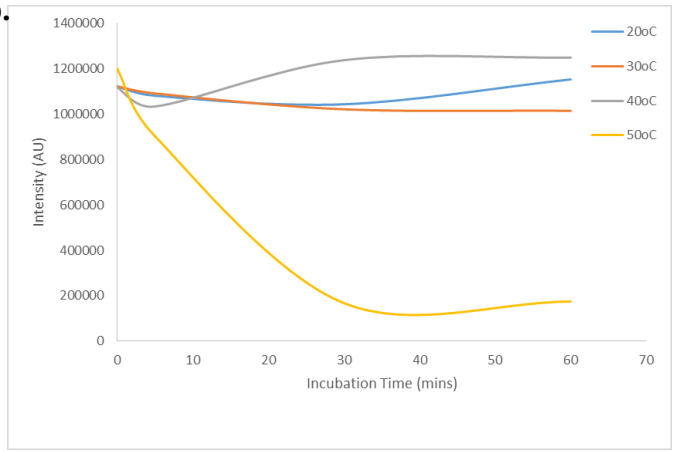
Figure 5.21. ΔV362 and ΔG363 bioluminescent activity measured at varying temperatures

.

a.



b.



Bioluminescent activity displayed by variants following incubation at various temperatures. a) ΔV362 b) ΔG363 assayed with 0.5µM of each enzyme, pre-incubated at different temperatures, added to 500µM LH_2 and 1mM ATP in chilled TEM buffer (pH 7.8) and spectra measured.

5.4. Further Discussion

The aim of this investigation was to characterise single amino acid deletion mutants identified in Chapter 4 to determine whether characteristics observed during 96-well format screening were true representations of the mutants and to conclusively determine whether single amino acid deletions in the loops selected enhance x11FLuc activity. As such single amino acid deletion mutants, ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363, derived from the Ω loop were purified and properties such as bioluminescence spectra, kinetics and specific activity were analysed.

With regards to determination of kinetic parameters involving the LH_2 substrate, whilst there was some alteration promoted due to deletions, these alterations were minimal, altering K_M by at most 5μM. However, with regards to the kinetic parameters involving the ATP substrate, larger changes were noted, whereby differences compared to x11FLuc amongst certain deletion variants were double that of the x11FLuc. It is of interest to note that with regards K_M, neighbouring deletions appeared to perform similarly, for example, the lowest K_M values for LH_2 identified were attributed to deletions ΔP359 and ΔG360 whilst the highest K_M values for ATP were attributed to deletions ΔA361, ΔV362, ΔG363, which displayed incremental increases of 147μM, 150μM and 188μM, respectively. It has been determined that the Ω loops are often involved in the function of the protein and molecular recognition (Fetrow *et al.*, 1995) and this functional role has been evidenced amongst a number of studies. The changes in K_M values for ATP suggest the Ω loop plays a role in the function of the protein and molecular recognition with regards to ATP and this is in-line with the literature (Johnson and Holyoak., 2012; Wiesner *et al.* 2010). In Chapter 3 molecular analysis identified the Ω loop within proximity to the ATP binding site which provides additional evidence that this may be the case.

With regards the spectrum of bioluminescence, all deletion mutants were similar to x11FLuc at optimum and physiological pH (7.4), which is pertinent to diagnostic assays with FLuc. Indeed, with regards to this characteristic, ΔP353, ΔP359, ΔG360, ΔA361, ΔV362 all showed no shift in □max or alteration in the FWHM. For ΔG363 on the other hand, the spectra were not stable, similar to x2FLuc. However, whilst acidic conditions promoted an increase in the FWHM and λmax of x2FLuc, alkali conditions

promoted such changes within ΔG363, whilst within acidic conditions it displayed no such changes.

The stability of the single amino acid deletion variants to pH and temperature was determined to identify whether deletions within the Ω loop region had retained the desirable characteristics of x11FLuc. Stability was determined by pH and thermal inactivation assays. Thermostability assays showed that whilst all deletions variants were unable to tolerate temperatures to the degree of x11FLuc, all outperformed x2FLuc. In addition to this, ΔP353, ΔP359, ΔG360, ΔA361, ΔV362, ΔG363 had either similar or lower pH tolerance in comparison to x11FLuc, and were superior to x2FLuc. This suggests that the Ω loop plays a role in catalytic activity, but does not play a vital role in dictating stability.

In Chapter 4, to determine which mutants to take forward to purification and characterisation, 96-well format assays were developed to isolate mutants with apparent enhancements with reference to kinetic parameters, specific activity and reduction in the K_M for the LH_2 substrate. Of the Ω loop variants, this strategy correctly identified that ΔP359 exhibited a reduced K_M for LH_2, whilst in addition to this it correctly identified that there was a trend throughout the Ω loop for enhancement of specific activity. In particular ΔG363 and ΔP359 has the highest specific activity of all mutants investigated which was correlated to the previous 96-well screening, lending weight to its applicability.

In this chapter we have determined that, following characterisation of pure protein, single amino acid deletions are useful in modulating characteristics of x11FLuc, including kinetics and bioluminescent spectra and SDD is useful technique that can be utilised to engineer FLucs. This adds to evidence that the employment of single amino acid deletional approaches, do not necessarily, nor typically disrupt the activity and can in fact enhance it. Therefore, FLuc protein engineering strategies should include such mutagenesis methods when aiming to improve existing proteins.

5.5. Conclusions

Here in this Chapter, single amino acid deletions occurring within the Ω loop, previously identified by *in vivo* colony screening and via a 96-well format as exhibiting potentially beneficial characteristics were purified and characterised with regards to kinetic parameters, specific activity and stability in terms of resistance to thermal inactivation and pH.

Following characterisation it has been identified that deletions occurring within the Ω loop confer wide ranging beneficial phenotypic alterations. ΔP359 and ΔG360 confer a lower K_M for the LH_2 substrate, whilst deletions of ΔA361, ΔV362 and ΔG363 confer an increase in the K_M for ATP. In addition this, 5 of 6 single amino acid deletions conferred a higher specific activity compared to x11FLuc, with ΔP359 and ΔG363 displaying the highest specific activity, markedly increased compared to the parental protein. Deletions within the Ω loop region do not retain the same thermostability as x11FLuc, however with regards to thermal and pH stability all mutants exhibit greater stability than the WTFLuc and x2FLuc, indicating that the 11 mutation of the x11FLuc continue to confer some additional stability.

Characterisations performed within this chapter also confirmed findings observed in Chapter 4, evidencing the utility of the 96-well format screens to identify important characteristics of mutants and the assay determined true representations of these characteristics.

These results add to growing evidence that the use of single amino acid deletions should not be overlooked as a method for protein engineering. This method improves the sampling of conformational space, which is not accessible to substitution methodologies and enables the discovery of novel mutants which confer a great number of beneficial phenotypic effects, such as kinetic alterations and changes in the emission colour of the protein.

Chapter 6

General Discussion

6.0. Chapter Summary

The aim of this project was to employ the use of single amino acid deletions within x11FLuc in order to address two questions i) to identify if, and to what extent, luciferase can tolerate single amino acid deletions in loop structures, and ii) to identify those deletions which result in desirable characteristics such as brightness, alteration of reaction kinetics, altered bioluminescence emission colours, pH tolerance and resistance to thermal inactivation.

6.1. The Utility and Novelty of Deletions Mutations in Protein Engineering

There is a misconception that amino acid deletions will likely compromise the structural integrity of a protein and abolish its function (Taylor *et al.,* 2004). Therefore, whilst protein engineering strategies such as rational or random substitution mutagenesis are heavily applied to identify altered or enhanced FLuc proteins, until now, the use of deletion mutagenesis has been completely overlooked in the field of FLuc engineering.

One of the key aims of this investigation was to deduce whether single amino acid deletions would be tolerated within the loop structures of FLuc, since there is growing stream of evidence that deletions may be useful as an alternative methodology to substitution mutagenesis for protein engineering (Simm *et al*., 1996; Arpino *et al.,* 2014). The rationale for targeting loops was based on previous studies whereby loops were determined to be most tolerant structural element (Taylor *et al.,* 2004; Arpino *et al*., 2014). Utilising the SDD mutagenesis strategy, amino acids within the loops M1-G10, L172- T191, T352-F368, D375-R387, D520-L526, K543-L550, were targeted for

sequential single amino acid deletion. From a total of 43 deletion mutants including both within regions close to the active site and within sites deemed to be more highly conserved, all bar two variants displayed bioluminescence. This indicated that the x11FLuc was able to tolerate deletions within its backbone to a greater extent than that noted in previous deletion models, such as GFP. Since FLuc is a more globular and flexible multi-domain protein compared to GFP (a rigid beta-barrel fold), it may tolerate better the structural rearrangements occurring as a result of the deletions.

In order to select mutants of interest from the 43 deletion variants, a 96-well format luminometric cell lysate assay was developed to estimate the specific activity of mutants, resistance to thermal inactivation, resistance to inhibition and to isolate mutants with kinetic alterations prior to purification and quantitative characterisation of selected proteins. This showed that sequential N-terminal deletions had little effect on the activity of x11FLuc, whilst deletions within L172- T191, D375-R387, D520-L526 and the C-terminal loop K543-L550 reduced overall activity. Not only were single amino acid deletions largely well-tolerated, through the course of the investigation, mutants were identified with enhanced properties in comparison to the x11FLuc template on which they were based. Deletions in the Ω loop T352-F368 caused increases in bioluminescence activity compared to native x11FLuc. Therefore, deletion mutants in the Ω loop were chosen for purification and biochemical characterisation. Of these, ΔA361, ΔV362 and to a greater extent ΔP359 and ΔG363 displayed higher specific activity. In addition to this, deletion mutants ΔP359 and ΔG360 displayed lower K_M for LH_2 whilst mutants ΔA361, ΔV362 and ΔG363 increased in their K_M for ATP. This highlighted the importance of sequentially deleting single amino acids along the length of each loop since neighboring deletion mutations did not always give rise to a similar characteristic. This also highlighted mutant ΔP359 as an enhanced version of x11FLuc for use in diagnostic assays carried out at temperatures up to 42^{o}C.

The bioluminescent spectra of deletion mutants within loops M1-G10, L172- T191, D375-R387 and K543-L550 had identical green emission spectra as native x11FLuc, whilst red shifts of up to 50nm were seen within some of the amino acid deletions occurring within loops T352-F368 and D520-L526. Therefore, deletions can promote enhancements with regards to activity and also colour. This may represent the incorporation of novel characteristics compared to substitution mutagenesis. It is

possible that this may include changes in the rates of successful FLuc folding, since the method may result in larger rearrangements of the backbone to improve the overall protein folding landscape. For example, deletion mutant ΔG363 bathochromically shifted under alkali conditions and yet resisted such changes under acidic conditions, the opposite to what is observed with WTFLuc or x2FLuc. Therefore, this work shows for the first time that that InDels may not only be tolerated but are beneficial, and can cause a broad range desirable phenotypes in FLuc.

To understand the role of deletions within protein engineering we may look to evolution. Within nature it is commonly accepted that evolution is largely as a result of small and gradual substitution mutations, and in the evolution of enzymes, such changes may result in changes in substrate specificity. However, they do not account for the larger transitions such as the creation of new protein folds. InDels bridge this gap since they can cause far larger perturbations of the structure (Arodz and Plonka, 2012). As such. Therefore, InDels are a prime example of mutations able to facilitate a wide range of changes and increase the possibility of generating novel protein functions.

6.2. The Role of the Ω Loop within Luciferase

Through the course of this investigation an Ω loop was identified within 10Å of the active site of FLuc (T352-F368) and was shown that deletions in this loop generated novel and useful mutants. The function, stability and folding of a protein is inherently linked to the structure of that protein. As mentioned, the Ω loop structure was first described in 1986 and since then the function of these structures has been elucidated (Fetrow *et al.*, 1995). Ω loops are non-regular secondary structures, whereby the characteristic shape is due to a large number of hydrogen bonds which are less periodic than seen in other structural elements. The less ordered structure of Ω loops has been shown to be important function, for example, substrate specificity or molecular recognition (Fetrow *et al* 1995). Ω loops may also be implicated in stability, therefore they may be classed as functional Ω loops, stability Ω loops, and folding Ω loops.

Interestingly, the phenotype of substitutions and insertions within this particular loop of FLuc have been reported in previous studies with particular focus on the thermostability

and colour variation observed when modifications are made at this site. E354R and D357Y are responsible for conferring resistance to thermal inactivation to x2FLuc (White *et al.,* 1996) and insertions within this region (see Chapter 3) increase the optimum temperature for bioluminescent activity and emission colour (Tafreshi *et al.,* 2007). In this Thesis, deletion mutants in the Ω loop exhibited similar or altered colours compared to x11FLuc, however certain of the single amino acid deletion mutants (ΔA361, ΔV362 and ΔG363) also caused an increase (doubling in some cases) in the K_M for ATP compared to x11FLuc. Therefore, the Ω loop may play a role in determining the binding affinity or adenylation with respect to ATP.

6.3. Alternative Screening Strategies

A 96-well format screens was used to reliably estimate characteristics of a large pool of mutants, such as resistance to thermal inactivation or inactivation by inhibitor compounds, which may be present in diagnostic assays in which FLuc is used. The screening methodology was optimised such that it enabled the estimation of such desirable characteristics without the need to purify the protein. This useful methodology correctly identified mutants with higher specific activity and a lower K_m for LH_2, however, the power of this technique for screening is the ability to determine characteristics of mutants that cannot be screened at the colony level, which can cause important mutant phenotypes to be overlooked. A clearly reproducible trend was identified in 96-well assays carried out at RT, which resembled a 'fingerprint' of activity caused by sequential deletions in loop regions spanning x11FLuc. For example, the N-terminal loop single amino acid deletions resulted in no significant changes in the activity of the luciferase, whereas deletion of 8 amino acids (together) is known to nearly abrogate activity (Wang et al., 2002; Sung and Kang, 1998). It is possible that the enhanced stability of x11FLuc compared to WTFLuc improves its tolerance to deletions. Single amino acid deletions within the loop L172- T191 caused a large reduction in activity whilst on the other hand, single amino acid deletion mutants within the Ω loop T352-F368 markedly increased activity in comparison to x11Fluc. However, deletions within the C-terminal loop (K543-L550) chosen in this study invariably and markedly reduced activity, as indicated by the consistent fingerprint pattern of activity in this region (Chapter 4). Conventionally those loops exhibiting lower activity such as

L172- T191 and the C-terminal would be discarded however utilsing the 96-well format screening strategy, loop L172- T191 was seen to be potentially important with regards to enzyme stability, since deletions made within this loop appeared to confer higher resistance to thermal inactivation. Within the remit of this study, the primary focus was to isolate mutants exhibiting enhancements in overall activity (mutants within the Ω loop), however, deletion mutations within the L172-T191 region may prove desirable for engineering higher thermostability of x11FLuc. This method addresses an additional issue raised by Kazlauskas and Bornscheuer (2010) who state that protein engineering successes (i.e. isolation of novel and useful phenotypes) should be accompanied by an experimentally supported hypothesis for the molecular basis of the novel function. It can be argued that all mutagenesis screens should not always discard non-functional proteins since exploration of their effects can lead to greater understand the mechanism of defined protein regions. However, colony screening methods are not easily able to accommodate or track the evolution of such variants, but this could be more easily achieved in 96-well or 384-well cell lysate formats as described in Chapter 4.

Using such comprehensive screening in combination with a rational or even a random deletion platform (such as MuDel – see Chapter 1), or in combination with rational or random substitutions and insertions, has the potential to derive many novel FLuc mutants in a simple and robust manner. Unlike other studies which do not screen libraries for more than one parameter, this investigation was able to sample many single amino acid deletions within loops and estimate multiple properties, which led to quick identification of important candidates, whilst in parallel estimating the effects that deletions may confer upon a structure-function of all mutants. The fingerprint patterns that emerged from the experiments highlight how specific regions may play a role in different characteristics of the protein, in terms of overall activity, colour, kinetics or stability.

6.4. Future Directions for Protein Engineering

Directed evolution is a powerful technique within protein engineering, whereby the characteristics of a protein can be selected based on the screening strategy employed. Directed evolution is limited by the diversity of the protein libraries that may be generated for screening since routine mutagenic strategies employ substitutions but do

not employ the use of insertion or deletion mutations, which increase the conformational space that may be sampled.

Deletions in proteins are engineered via two primary existing methods: firstly, a targeted approach based on primer design (as conducted in this study) and directed evolution by novel methods, for example, a technique which allows for random large scale introduction of single amino acid deletions throughout an entire protein, is known as Trinucleotide exchange (Tri-NEX) (Figure 6.1.).

TriNEx utilses a transposon based approach, using a transposon termed MuDel as derived from bacteriophage Mu and has been can be engineered to carry useful genes. A transposon is a linear strand of DNA which can insert into DNA, due to "Transposase Recognition Elements" (TREs) which lie within the terminal ends of the transposon which are identified by a MuA transpose through mediate the insertion of the transposon into DNA. During insertion, DNA cleavage is normally staggered due to the position of hydroxyl groups, resulting in duplication at the site of the insertion. In the case of the Mu/ MuA system, there is a 5 base pair (bp) stagger resulting in 5bp site duplication flanking the transposon after insertion.

TriNEx utilises the Mu/ MuA transposition mechanism to introduce trinucleotide deletion mutations translating into deletions of single amino acids. This technique is useful as the libraries generated with this technique are random and contain many different trinculetoide deletions. The transposon has been engineered to include a marker for selection by chloramphenicol resistance and to contain *MlyI* restriction sites 1bp from the terminals of the transposon. These restriction sites allow for the removal of the transposon in combination with a removal of a trinucleotide from the gene of interest.

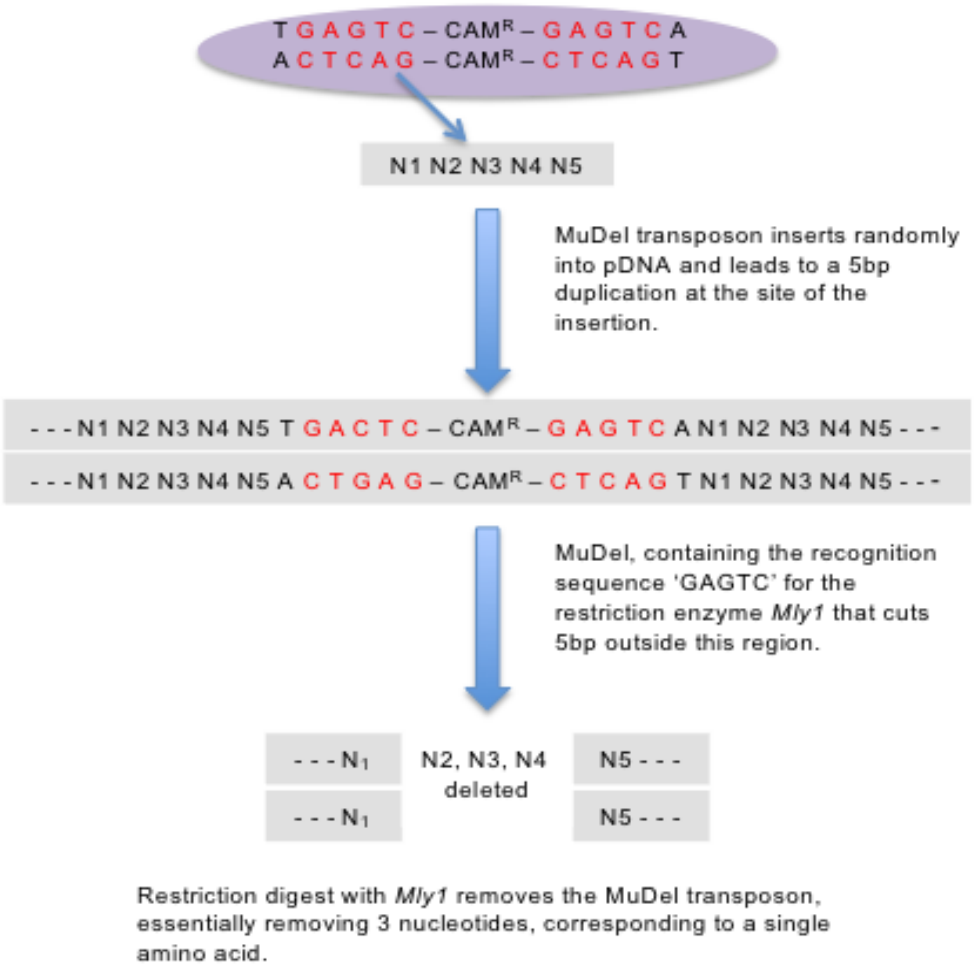
TriNEx is not simply limited to the removal of single amino acid deletions since out of frame deletions promote a deletion and give rise to a subsequent neighboring substitution mutation increasing the number of potential mutants available for screening. The ability to derive deletion and substitution mutations is highly desirable since successful protein engineering strategies often mirror that situation in nature.

The accommodation of InDels within natural proteins tends to be dependent upon the presence of substitution in the spatial vicinity of the accepted InDel to enable the effect of the deletion. This being the case, future protein engineering strategies should incorporate both substitution methodologies in conjunction within deletions.

Whilst this work is confined to the use of single amino acid deletions within loops within luciferase due to the more so tolerant nature of loops, previous studies within eGFP highlighted that a deletion within an α-helix allowed a beneficial registry shift resulting in a new polar interaction network. It may be considered that whilst deletions within loops are well tolerated within x11FLuc and have indeed conferred desirable properties onto the protein, deletions within secondary structural elements may hold a repertoire of many more novel proteins, which promote new interactions in the enzyme. It is desirable that future work investigates the use of single amino acid deletions within these secondary structural elements to not only determine tolerance within these structures but to identify whether beneficial registry shifts may be imparted as a result.

Future work should not only involve the investigation of deletions within the secondary elements of x11FLuc, but should aim to elucidate the underlying causality for the improved function of the deletion mutants obtained within this study. To determine this, crystal structures may be determined for both x11FLuc and the deletion mutants, paying specific attention to those of particular interest, such as ΔP359 and ΔG363.

Figure 6.1. Schematic of TRiNEx



Schematic of the mechanism of TRiNEx in order to generate a library of single amino acid deletion mutants.

References

Ai H, Shaner NC, Cheng Z, Tsein RY and Campbell RE (2007). *Exploration of New Chromophore Structures Leads to the Identification of Improved Blue Fluorescent Proteins.* Biochemistry 46: 5904-5910.

Airth RL, Rhodes WC and McElroy WD (1958). *The function of coenzyme A in luminescence.* Biochem. Biophys. Acta 27: 519-532.

Ando Y., Niwa K, Yamada N, Enomoto T, Irie T, Kubota H, Ohmiya Y and Akiyama H (2008). *Firefly bioluminescence quantum yield and colour change by pH sensitive green emission.* Nature Photonics 2:44-47.

Arpino JAJ, Reddington SC, Halliwell LM, Rizkallah PJ and Jones DD (2014). *Random Single Amino Acid Deletion Sampling Unveils Structural Tolerance and the Benefits of Helical Registry Shift on GFP Folding and Structure.* 22:889-898.

Baldisseri, D., Torchia, D., Poole, L. and Gerlt, J (1991). *Deletion of the Omega-Loop in the Active Site of Staphylococcal Nuclease*. 2. Effects on Protein Structure and Dynamics. Biochemistry. 30, 3628-3633.

Baldwin AJ, Busse K, Simm AM and Jones, D.D. (2008). *Expanded Molecular Diversity Generation during Directed Evolution by Trinucleotide Exchange (TriNex).* Nucleic Acids Research. 13: e77.

Bradford MM (1976). *A rapid and sensitive method for the quantification of microgram quantities of protein utilising the principle of protein dye binding.* Anal. Biochem. 72: 284-254.

Branchini BR, Magyar RA, Murtiashaw MH, Anderson SM and Zimmer M (1998). *Site-directed mutagenesis of Histidine 245 in firefly Luciferase: A Proposed Model of the Active Site*. Biochemistry 37: 15311-15319.

Branchini BR, Murtiashaw MH, Magyar RA and Anderson SM (2000). *The role of lysine 529, a conserved residue of the acyl-adenylate-forming enzyme superfamily, in firefly luciferase.* Biochemistry 39: 5433-40.

Branchini BR, Magyar RA, Murtiashaw MH and Portier NC (2001). *The role of active site residue arginine 218 in the firefly luciferase bioluminescence.* Biochemistry 40: 2410-8.

Branchini BR, Murtiashaw MH, Magyar RA, Portier NC, Ruggiero MC and Stroh JG (2002). *Yellow-green and red firefly bioluminescence from 5,5-dimethyloxyluciferin.*J. Am. Chem. Soc. 124: 2112-3.

Branchini BR, Southworth TL, Murtiashaw MH, Boije H and Fleet SE (2003). *A mutagenesis study of the putative luciferin binding site residues of firefly luciferase.* Biochemistry 42: 10429-36.

Branchini BR, Southworth TL, Murtiashaw MH, Magyar RA, Gonzalez SA, Ruggiero MC and Stroh JG (2004). *An alternative mechanism of bioluminescence colour determination in firefly luciferase.* Biochemistry 43: 7255-62.

Branchini BR, Southworth T, Khattak NF, Michelini E and Roda A (2005). *Red- and green-emitting firefly luciferase mutants for bioluminescent reporter applications.* Anal. Biochem. 345: 140-8.

Branchini BR., Ablamsky DM, Davis AL, Southworth TL, Butler B, Fan F. Jathoul AP and Pule MA (2009). *Red-emitting luciferases for bioluminescence reporter and imaging applications.* Anal. Biochem. 396, 290-297.

Branchini BR, Southworth TL, Fontaine DM, Davis AL, Behney CE and Murtiashaw MH (2014). *A Photinus pyralis and Luciola italic chimeric firefly luciferase produced enhanced bioluminescence.* Biochemistry 53: 6287-9 25.

Branchini BR, Southworth TL, Fontaine DM, Kohrt D, Talukder M, Michelini E, Cevenini L, Roda A and Grossel MJ (2015). *An enhanced chimeric firefly luciferase-inspired enzyme for ATP detection and bioluminescence reporter and imaging applications*. Anal Biochem 484: 148-53.

Branchini BR, Behney CE, Southworth TL, Fontaine DM, Gulick AM, Vinyard DJ and Brudvig GW (2015*). Experimental Support for a Single Electron-Transfer Oxidation Mechanism in Firefly Bioluminescence*. J Am Chem Soc. 137: 7592-5.

Brannigan J and Wilkinson A (2002). *Protein Engineering 20 Years On*. Nature Reviews Molecular Cell Biology. 3: 964-970.

Brovko LY, Gandleman OA, Polenova TE and Ugarova NN (1994). *Kinetics of bioluminescence in the firefly luciferin-luciferase system.* Biochem. (Moscow) 59: 195-201.

Conti E, Franks NP and Brick P (1996). *Crystal structure of firefly luciferase throws light on a superfamily of adenylate-forming enzymes.* Structure 4: 287-98.

Day JC, Tisi LC and Bailey MJ (2004). *Evolution of beetle bioluminescence: the origin of beetle luciferin.* Luminescence 19: 8-20.

DeLuca M and McElroy WD (1974). *Kinetics of the firefly luciferase catalysed reactions.* Biochem. 13: 921-5.

DeLuca M (1976). Firefly luciferase. In Meister, A. (Ed.). Advances in enzymology and Related Areas of Molecular Biology, 44, 37-68, John Wiley and Sons, New York.

DeLuca M and McElroy WD (1978). *Purification and Properties of firefly luciferase.* Methods in Enzymology 57: 3-15.

Dubuisson M, Marchand C and Rees JF (2004). *Fire fly luciferin as antioxidant and light emitter: the evolution of insect bioluminescence.* Luminescence 19: 339-44.

Dunker K, Brown C, Lawson D, Iakoucheva LM and Obradovic Z (2002). *Intrinsic Disorder and Protein Function.* Biochemistry 41: 6573-82.

Eisner T, Wiemer DF, Haynes LW, Meinwald J (1978). Lucibufagins: Defensive steroids from the fireflies Photinus ignitus and P. marginellus (Coleoptera: Lampyridae). Proc. Natl. Acad. Sci 75: 905-8.

Fetrow JS (1995). *Omega loop: non regular secondary structure significant in protein function and stability.* FASEB J. 9: 708-17.

Fontes R, Dukhovich A, Sillero A, Sillero MA (1997). *Synthesis of dehydroluciferin by firefly luciferase: effect of dehydroluciferin, coenzyme A and nucleoside triphosphates on the luminescent reaction.* Biochem Biophys Res Commun 237: 445-50.

Fontes R, Ortiz B, Diego A, Sillero A and Sillero MAG (1998). *Dehydroluciferyl-AMP is the main intermediate in the luciferin dependent synthesis of Ap4A catalyzed by firefly luciferase*. FEBS Lett 438: 190-194.

Ford S, Buck L and Leach F (1995). *Does the Sulfhydryl or the Adenine Moiety of CoA Enhance Firefly Luciferase Activity?* Biochim. Biophys. Acta. 1252: 180-4.

Fraga H, Fernandes D, Fontes R, Esteves da Silva JC (2005). *Coenzyme A affects firefly luciferase luminescence because it acts as a substrate and not as an allosteric effector*. FEBS J 272: 5206-16.

Gandleman OA, Church VL, Moore CA, Kiddle G, Carne CA, Parmar S, Jaal H, Tisi LC and Murray JAH (2010). *Novel bioluminescent quantitative detection of nucleic acid amplification in real time.* PloS One 5:

Gould ST, Keller G and Subramani D (1987). *Identification of peroxisomal targeting signcals located at the carboxy terminus of four peroxisomal proteins* J Cell Biol 107: 897-905.

Green AA and McElroy WD (1956). *Crystalline firefly luciferase*. Biochem. Biophys. Acta 20: 170-176.

Gulick AM (2009). *Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase*. ACS Chem Biol. 4, 811–827.

Hall MP, Gruber MG, Hannah RR, Jennens-Clough ML and Wood KV (1999). *stabilisation of firefly luciferase using directed evolution.* In *Bioluminescence and chemiluminescence: perspectives for the 21st centery*. Editors: Roda A, Pazzagli M, Kricka LJ and Stanley PE. Wiley, Chichester. Pgs 392-5.

Harvey EN (1940). *Living Light*. Hafner, New York.

Haddock SHD, Moline MA and Case JF (2010). *Bioluminescence in the Sea.* Annual Review of Marine Science 2: 443-493.

Hastings JW and Tu SC (1981). Bioluminescence of Bacterial Luciferase. Annals of the New York Academy of Sciences 366: 315-327.

Ishida T and Kinoshita K (2007). *PrDOS: prediction of disordered protein regions from aminoacid sequence.* Nucleic Acids Res.

Iwano S, Obata R, Miura C, Kiyama M, Hama K, Nakamura M, Amano Y, Kojima S, Hirano T, Maki S and Niaw H (2013). *Development of simple firefly luciferin analogs emitting blue, green, read and near-infrared biological window light.* Tetrahedon 69: 3847-3856.

Jawhara s and Mordon S (2004). *In vivo imaging of bioluminescent Echerichia coli in a cutaneous wound infection model for evaluation of an antibiotic therapy.* Antimicrob. Agents. Chemother. 48: 3436-41.

Jathoul A (2008). Activity of Firefly Luicferase with 6'- Amino-D-Luciferin. Thesis

Jathoul A, Law E, Gandleman O, Pule M, Tisi L and Murray JAH (2012). *Development of a pH-Tolerant Thermostable Photinus pyralis luciferase for brighter in vivo imaging.* Bioluminescence – Recent Advances in Oceanic Measurements and Laboratory Applications.

Jathoul AP, Grounds H, Anderson J C and Pule MA (2014). *A Dual-Color Far-Red to Near-Infrared Firefly Luciferin Analogue Designed for Multiparametric Bioluminescence Imaging*. Angew. Chem. Int. Ed. 53: 13059–13063.

Koksharov MI and Ugarova NN (2008). *Random mutagenesis of Luciola mingrelica firefly luciferase. Mutant enzymes with bioluminescence spectra showing low pH sensitivity.* Biochemistry 73: 862-9.

Koo JY, Schmidt SP and Schuster GB (1978). *Bioluminescence of the firefly: key steps of the formation of the electronically excited state for model systems*. PNAS 75: 30-3.

Kepler TB, Liao HX, Alam SM, Bhaskarabhatla R, Zhang R, Yandava C, Stewart S, Anasti K, Kelsoe G, Parks R, Lloyd KE, Stolarchuk C, Pritchett J, Solomon E, Friberg E, Morris L, Karim SS, Cohen MS, Walter E, Moody MA, Wu X, Altae-Tran HR, Georgiev IS, Kwong PD, Boyd SD, Fire AZ, Mascola JR and Haynes BF. *Immunoglobulin gene insertions and deletions in the affinity maturation of HIV-1 broadly reactive neutralizing antibodies.* Cell Host Microbe 16: 304-13.

Law GHE, Gandleman OA, Tisi LC, Lowe CR and Murray JAH (2002). *Altering the surface hydrophobicity of firefly luciferase.* In *Bioluminescence and Chemiluminescence: Progress and current applications.* Editors: Stanley PE and Kricka LJ. World Scientific, Singapore. Pgs 189-92.

Law GHE, Gandleman OA, Tisi LC, Lowe CR and Murray JAH (2006). *Mutagenesis of solvent exposed amino acids in Photinus pyralis luciferase improves thermostability and pH tolerance.* Biochem. J. 397: 305-312.

Lemasters JJ and Hackenbrock CR (1977). *Kinetics of the product inhibition during firefly luciferase luminescence.* Biochem 16: 445-7.

Lutz S. and Patrick W M (2004) *Novel Methods for directed evolution of enzymes: quality, not quantity*. Curr Opin Biotechnol. 15, 291-297.

Maguire CA, Deliolanis NC, Pike L, Niers JM, Tjon-Kon-Fat LA, Sena-Esteves M and Tannous BA (2009). *Gaussia luciferase variant for high-throughput functional screening applications*. Analytical Chemistry 81: 7102-7106.

Marques SM and Esteves da Silva JC (2009). *Firefly bioluminescence: a mechanistic approach of luciferase catalysed reactions*. IUBMB Life 61: 6-17.

McElroy WD and Hastings JW (1955). *Biochemistry of firefly luminescence.* In *the luminescence of biological systems*. Editor: Johnson FH. A.A.A.S. Press, Washington DC, USA. Pgs. 161-98.

McElroy WD and Seliger HH (1961). *Mechanisms of bioluminescent reactions.* In McElroy WD and Glass B. Light and Life. The John Hopkins University Press, Baltimore. Pgs 219-257.

McElroy WD and Seliger HH (1962). *Mechanism of action of firefly bioluminescence.* Fed. Proc. Fed. Am. SOc. Exp. Biol 21: 2006-1012.

Moradi A, Hosseinkhani S, Naderi-Manesh H, Sadeghizadeh M and Alipour B (2009). *Effect of charge distribution in a flexible loop on the bioluminescence colour of firefly luciferases.* 48: 575-582.

Nakajima Y, Kobayashi K, Yamagishi K, Enomoto T, Ohmiya Y (2004). *cDNA cloning and characterisation of a secreted luciferase from the luminous Japanese ostracod, Cypridina noctiluca.* Biosci Biotechnol Biochem 68: 565-70.

Nakamura M, Makia S, Amanoa Y, Ohkitaa Y, Niwab K, Hiranoa T, Ohmiyab Y and Niwaa H (2005). *Firefly luciferase exhibits bimodal action depending on the luciferin chirality*. Biochemical and Biophysical Research Communications 331:471-475.

Nakamura M, Niwa K, Maki S, Hirano T, Ohmiya and Niwa H (2006). *Construction of a new firefly bioluminescence system using L-luciferin as substrate.* Tetrahedon Lett 47:1197-1200.

Nakatani N, Hasegawa J and Nakatsuji H (2007). *Red light chemiluminescence and yellow-green light in bioluminescence: colour tuning mechanism of firefly, Photinus pyralis, studied by the symmetry adapted cluster configuration interaction method.* JACS 129: 8756-65.

Niwa K, Nakamura M and Ohmiya Y (2006). *Steroisomeric bio-inversion key to biosynthesis of firefly D-luciferin.* FEBS Letts 580: 5283-5287.
Niwa K, Ichino Y, Kumata S, Nakajima Y, Kato D, Viviani VR and Ohmiya Y (2010). *Quantum yields and kinetics of the firefly bioluminescence reaction of beetle luciferases.* Photochem Photobiol 86: 1046-9.

Neylon C (2004). *Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution.* Nucleic Acids Res. 32, 1448-1459.

Novagen (2011). *pET System Manual 11th Edition.*

Oba Y, Ojika M and Inouye S (2003). *Firefly luciferase is a bifunctional enzyme: ATP-dependent monooxygenase and a long chain fatty acyl-CoA synthetase*. FEBS Lett 540: 251-4.

Oba Y, Sato M and Inouye S (2006). *Cloning and characterisation of the homologous genes of firefly luciferase in the mealworm beetle, Tenebrio molitor*. Insect Mol Biol 15: 293-9.

Oldfield CJ, Ulrich EL, Cheng Y, Dunker AK and Markley JL (2005). *Addressing the intrinsic disorder bottleneck in structural proteomics*. Proteins. 2005; 59:444–453.

Osuna J, Perez-Blancas A and Soberon X (2002). *Improving a circularly permuted TEM-1 beta-lactamase by directed evolution.* Protein Eng. 15: 463-70.

Prebble SE, Price RL, Lingard B, Tisi LC and White PJ (2001). *Protein engineering and molecular modelling of firefly luciferase.* In Proc 11th Int. Symp. Biolum. Chemilum. Editors: Case JF, Herring PJ, Robinson BH, Haddock SHD, Kricka LJ abd Stanley PE. World Scientific, Singapore. Pgs 181-4.

Price, S. and Nagai, K. (2002). *Protein Engineering as a Tool for Crystallography*. Current Opinion in Biotechnology. 6: 425-430.

Purtov KV, Petushkov VN, Baranov MS, Mineev KS, Rodionova NS, Kaskova ZM, Tsarkova AS, Petunin AI, Bondar VS, Rodicheva EK, Medvedeva SE, Oba Y, Arseniev AS, Lukyanov S, Gitelson JI and Yampolsky IV (2015). The Chemical Basis of Fungal Bioluminescence. Angew. Chem. 127: 8113

Rath A, Glibowicka M, Nadeau VG, Chen G and Deber CM (2008). *Detergent binding explains anomalous SDS-PAGE migration of membrane proteins*. PNAS. 106: 1760-65

Rice BW and Contag CH (2009). *The importance of being red.* Nature Biotechnology 27: 624-625.

Ronaghi M. (2001). *Pyrosequencing sheds light on DNA sequencing*. *Genome Res.* 11: 3-11.

Sala-Newby GB and Campbell AK (1994). *Stepwise removal of the C-terminal 12 amino acids of firefly luciferase results in graded loss of activity*. Biochem. Biophys. Acta. 1206: 155-60.

Seliger HH and McElroy WD (1959). *Quantum yield in the oxidation of firefly luciferin.* Biochem. Biophys. Res. Comm 1: 21-24.

Seliger HH, McElroy WD, White EH and Field GF (1961). *Stereospecificity and firefly bioluminescence, a comparison of natural and synthetic luciferin.* Proc. Natl. Acad. Sci. 47: 1129-1134

Simm A, Baldwin A, Busse K and Jones D (2007). *Investigating the Protein Structural Plasticity by Surveying the Consequence of an Amino Acid Deletion from TEM-1 β-lactamase*. Federation of European Biochemical Societies. 581, 3904-3908.

Sherf BA and Wood KV (1994). *Firefly luciferase engineered for improved genetic reporting.* Promega Notes 49: 14-21.

Shimomura O and Johnson FH (1975). *Regeneration of the photoprotein aequorin.* Nature 256: 236-238.

Shimomura O (1995). *A short story of Aequorin.* Biol. Bull 189: 1-5.

Shortle D and Sondek J (1995). *The Emerging Role of Insertions and Deletions in Protein Engineering*. Current Opinion in Biotechnology. 4: 387-393.

Shrope M (2007). *Marine biology: Lights in the deep.* Nature. 450: 472-474

Sun YQ, Liu J, Wang P, Zhang J and Guo W (2012). *D-luciferin analogues: a multicolor toolbox for bioluminescence imaging*. Angew. Chem. Intl. Ed. 51: 8438-30.

Sundlov JA, Fontaine DM, Southworth TL and Branchini BR (2012). *Crystal Structure of Firefly Luciferase in a Second Catalytic Conformation Supports a Domain Alternation Mechanism.* Biochemistry 51: 6493-6495.

Sung D and Kang H (1998). *The N-terminal amino acid sequences of the firefly luciferase are important for stability of the enzyme.* Photochem. Photobiol 68: 749-53.

Sillero A and Sillero MAG (2000). *Synthesis of dinucleoside polyphosphates catalyzed by firefly luciferase and several ligases*. Pharmacol. Ther 87: 91-102.

Svetlov MS, Kolb VA and Spirin A.S. (2007). *Folding of the firefly luciferase polypeptide chain with immobilised C- terminus*. Mol. Biol. 41: 96-102.

Taurianen S, Virta M Chang W and Karp M (2000). *Measurement of firefly luciferase reporter gene activity from cells and lysates using Escherica coli arsenate and mercury sensors*. Anal. Biochem. 272, 191-198.

Taylor, M.S., Ponting, C.P. and Copley, R.R. (2004). *Occurrence and Consequences of Coding Sequence Insertions and Deletions in Mammalian genomes*. Genome Research. 14, 555–566.

Tafreshi NK, Hosseinkhani S, Sadeghizadeh M, Sadeghi M, Ranjbar B and Naderi-Manesh H (2007). *The influence of insertion of a critical residue (Arg356) in structure and*

bioluminescence spectra of firefly luciferase. J. Biol. Chem. 282: 8641-7.

Telford WG, Moss MW, Morseman JP and Allnutt FC (2001). *Cyanobacterial stabilized phycobilisomes as fluorochromes for extracellular antigen detection by flow cytometry*. J Immunol

Thompson EM, Nagata S and Tsuji FI (1991). *Vargula hilgendorfii luciferase: a secreted reporter enzyme for monitoring gene expression in mammalian cells*. Gene 96: 257-262.

Thorne N, Inglese J and Auld DS. (2010). *Illuminating insights into firefly luciferase and other bioluminescent reporters used in chemical biology*. Chem Biol.

Tisi LC, Lowe CR and Murray JAH (2001). *Mutagenesis of solvent exposed hydrophobic residues in firefly luciferase.* In Proc. Proc 11th Int. Symp. Biolum. Chemilum. Editors: Case JF, Herring PJ, Robinson BH, Haddock SHD, Kricka LJ abd Stanley PE. World

Tisi LC, Law EL, Gandleman O, Lowe CR and Murray JAH (2002). *The basis of the bathochromic shift in the luciferase from Photinus pyralis.* In Bioluminescence and Chemiluminescence: Progress and Current Applications edited by Stanley PE and Kricka LJ. Pgs 57-60. Word Scientific, Singapore.

Tisi LC, White PJ, Squirrel DJ, Murphy MJ, Lowe CR and Murray JAH (2002b). *Development of a thermostable firefly luciferase.* Anal. Chim. Acta 457: 115-23 Scientific, Singapore. Pgs 189-92.

Tsien, R.Y. (1998). *The green fluorescent protein*. Annu. Rev. Biochem 67: 509-544.
Tu SC and Mager HIX (2008). *Biochemistry of Bacterial Bioluminescence.* Photochemistry and Photobiology 62: 615-624.

Ugarova (1989). *Luciferase of Luciola mingrelica fireflies. Kinetics and regulation*

mechanism. J.Biolum. Chemilum. 4:406-18.

Ugarova NN and Brovko LY (2002). *Protein structure and bioluminescent spectra for firefly bioluminescence.* Luminescence 17: 321-30.

Ugarova NN (2008). *Interaction of firefly luciferase with substrates and their analogs: a study using fluorescence spectroscopy methods.* Photochem. Photobiol. Sci. 7: 218-27.

Vassel N, Cox C, Naseem R, Morse V, Evans R, Power R, Brancale A, Wann K and Campbell A (2012). *Enzymatic activity of albumin shown by coelenterazine chemiluminescence*. Luminescence 27: 234–241

Viviani VR, Uchida A, Viviani W and Ohmiya Y (2002). *The influence of Ala243 (Gly247), Arg215 and Thr226 (Asn230) on the bioluminescence spectra and pH sensitivity of railroad worm, click beetle and firefly luciferases.* Photochem. Photobiol 76: 538-44.

Wang XC, Yang J, Huang W, He L, Yu JT, Lin QS, Li W and Zhou HM (2002). *Effects of removal of the N terminal amino acid residues on the activity and conformation of firefly luciferase*. Int. J., Biochem. Cell. Biol. 34: 983-91.

White EH, McCapra F, Field GF and McElroy WD (1961). *The structure and synthesis of firefly luciferin*. J. Am. Chem. Soc 83: 2402-2403.

White EH, Worther H, Seliger and McElroy WD (1966). *Amino analogs of firefly luciferin and activity thereof.* JACS 88: 2015-8.

White EH, Rapport E, Seilger HH and Hopkins TA (1971). *The chemi- and bioluminescence of firefly luciferin: An efficient chemical production of electronically excited states.* Bioorg. Chem 1:92-122.

White EH, Steinmetz MG, Miano JD, Wildes PD and Morland R (1980). *Chemi- and*

bioluminescence of firefly luciferin. JACS 102: 3199-3208.

White PJ, Squirrell DJ, Arnaud P, Lowe CR and Murray JAH (1996). *Improved thermostability of the North American firefly luciferase: saturation mutagnesis at position 354. Biochem. J.* 319: 343-50.

Wright PE and Dyson HJ (1999). *Instrinically unstructured proteins: re-assessing the protein structure-function paradigm.* J. Mol. Biol. 293: 321-31.

Wilson T and Hastings JW (1998). *Bioluminescence*. Annu Rev Cell Dev Biol 14: 197-230.

Yaginuma H, Kawai S, Tabata KV, Tomiyama K, Kakizuka A, Komatsuzaki T, Noji H and Wood KV and DeLuca M (1987). *Photographic detection of luminescence in Eschericheria coli containing the gene for firefly luciferase.* Anal. Biochem. 161: 501-507.

Imamura H (2014). *Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging.* Scientific Reports 4

Zako T, Ayabe T, Aburatani N, Kamiya A, Kitayama H, Ueda T and Nagamune (2003). *Luminescent and substrate binding activities of firefly luciferase N-terminal domain*. Biochem. Biophys. Acta. 1649: 183-189.

Zhang Z and Gerstein M (2003). *Patterns of nucleotide substitution, insertion and deletion in the human genome inferred from pseudogenes*. Nucleic Acids Res. 31: 5338-48.